

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

M. MASUDA et al.
1/3/01
Q62568
10f1

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application:

2000年 1月 11日

出願番号
Application Number:

特願2000-003041

出願人
Applicant(s):

日本電気株式会社

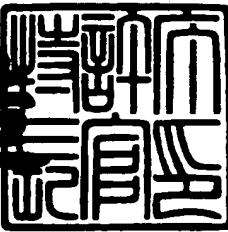
JC912 U.S. PRO
09/752520
01/03/01



CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年12月 8日

特許庁長官
Commissioner,
Patent Office

及川耕三


出証番号 出証特2000-3101718

【書類名】 特許願

【整理番号】 49210398

【提出日】 平成12年 1月11日

【あて先】 特許庁長官殿

【国際特許分類】 H04L 12/56

【発明者】

【住所又は居所】 東京都港区芝五丁目 7番1号 日本電気株式会社内

【氏名】 升田 道雄

【発明者】

【住所又は居所】 東京都港区芝五丁目 7番1号 日本電気株式会社内

【氏名】 有川 寿秋

【発明者】

【住所又は居所】 東京都港区芝五丁目 7番1号 日本電気株式会社内

【氏名】 山田 憲晋

【特許出願人】

【識別番号】 000004237

【氏名又は名称】 日本電気株式会社

【代理人】

【識別番号】 100065385

【弁理士】

【氏名又は名称】 山下 穏平

【電話番号】 03-3431-1831

【手数料の表示】

【予納台帳番号】 010700

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

特2000-003041

【包括委任状番号】 9001713

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 マルチレイヤクラス識別通信装置と通信装置

【特許請求の範囲】

【請求項1】 入力インターフェースとスイッチ回路と出力インターフェースとを備え、通信回線の複数の通信品質のレベルを選択可能なマルチレイヤクラス識別通信装置において、

前記通信回線に接続した入力インターフェースに、受信したIPパケットのヘッダ情報(IPヘッダに代表されるレイヤ3情報、およびその上位レイヤに相当するTCP/UDPヘッダに代表されるレイヤ4情報の組合せ)から前記IPパケット内におけるクラス識別子を解決し、個々のIPパケット処理を前記入力インターフェース内で取り扱うサービス品質(Internet Protocol Quality of Service: IP-QoS)コードを割当てる割当手段を有して、前記スイッチ回路で宛先アドレスに向けて前記通信回線を選択し、前記出力インターフェースからレイヤ4以上の上位レイヤ毎に出力スケジューリングを実施することを特徴とするマルチレイヤクラス識別通信装置。

【請求項2】 上記IPコードの割当て(トラヒッククラスのmapping)に關し、前記IPパケットのIPヘッダおよび、TCPヘッダ内の複数フィールドの任意の組み合わせで優先トラヒックを規定できる手段を有することを特徴とする請求項1に記載のマルチレイヤクラス識別通信装置。

【請求項3】 送信側のクラス別スケジューリング方式として、WRR(Weighted Round Robin Scheduling)と、固定優先スケジューリングを共存させ、前記クラス識別子による各クラスを固定優先で選択できるとともに、最小帯域指定も可能となる手段を有することを特徴とする請求項1に記載のマルチレイヤクラス識別通信装置。

【請求項4】 前記装置は、前記入力インターフェースと、出力インターフェースと、前記入力インターフェースと前記出力インターフェースとの間に宛先アドレスに向けて切り換えを行う切り換えスイッチと、前記入力インターフェースと前記出力インターフェースとの間に前記切り換えスイッチの切り換えタイミングを指示するスケジューラとを備え、前記入力インターフェースおよび前記出力

インターフェース、および前記切り替えスイッチが連動して動作することにより、前記IP-QOSコードにより、前記スケジューラの指示により所定の優先制御を行う有線制御手段を有することを特徴とする請求項1に記載のマルチレイヤクラス識別通信装置。

【請求項5】 前記IP-QOSコード単位に統計多重効果を得るために、複数のIPフローパケットを集約して共有できるキュー管理を行うキュー管理手段を有することを特徴とする請求項1に記載のマルチレイヤクラス識別通信装置。

【請求項6】 前記入力インターフェース、および前記出力インターフェースに、前記IP-QOSコード単位にトラヒックを監視し、過剰なトラヒックの規制を行うトラヒック規制手段を有することを特徴とする請求項4に記載のマルチレイヤクラス識別通信装置。

【請求項7】 前記クラス識別子は、EF (Expedited Forwarding (Premium service)) クラスと、AF (Assured Forwarding Service) クラスと、BE (Best Effort Service) クラスの3種類のトラヒック (サービスクラス) であることを特徴とする請求項4に記載のマルチレイヤクラス識別通信装置。

【請求項8】 通信回線の複数の通信品質のレベルを選択可能な入力インターフェースと出力インターフェースとを備えた通信装置において、

前記入力インターフェースは、入力するIPパケットからIPパケットヘッダ情報とTCPヘッダ情報を抽出するIPパケット受信部と、前記IPパケットヘッダ情報を検索キーとしてIP-QOSコード (クラス識別子) 解決メモリをアクセスして解決対象となるクラス識別子を取得するIP-QOSクラス決定部と、前記IPパケット受信部より受信したIPパケットデータと前記IP-QOSクラス決定部において解決した該IPパケットに対応するIP-QOSクラスコードにより宛先特定の優先制御を行う受信側スイッチインターフェース制御部と、前記優先制御機能と送信要求を前記クラス識別子単位に前記出力インターフェースカードに送出する受信側スイッチインターフェースとを備え、

前記IP-QOSクラス決定部は、前記IP-QOSクラス毎に設定された転送可能容量を超える過剰なトラヒックの流入を監視し、転送可能容量を超える場

合は、該当するIPパケットを廃棄、あるいは、前記送信優先度を下げるポリシング（監視）制御を行う手段を有していることを特徴とする通信装置。

【請求項9】 通信回線の複数の通信品質のレベルを選択可能な入力インターフェースと出力インターフェースとを備えた通信装置において、

前記入力インターフェースから送出された優先制御されたIPパケットを受信する出力インターフェースは、受信した前記IPパケットを送信側ペイロードメモリに格納すると同時に該IPパケット情報を生成し FIFOメモリに書き込む送信側スイッチインターフェースと、前記IPパケット情報内に収容されたIP-QoSクラスコードに基づいて、クラス毎にキューイングを行い優先的に送信要求し且つWRR (Weighted Round Robin) 方式に基づいたスケジューリング機能により要求サービス品質を提供する制御機能を有するIP-QoSクラス別スケジューラと、前記送信側スイッチインターフェースからのIPパケットを前記IP-QoSクラス別スケジューラの優先度でスケジューリングされた順序に従って送出する送信側スイッチインターフェース制御部と、前記IPパケットを下位層のデータリンク層及びネットワークアクセス層に送出するIPパケット送信部とからなることを特徴とする通信装置。

【請求項10】 請求項8又は9に記載された通信装置において、前記入力インターフェースに、受信した前記IPパケットのヘッダ情報(IPヘッダに代表されるレイヤ3情報、およびその上位レイヤに相当するTCP/UDPヘッダに代表されるレイヤ4情報の組合せ)から前記IPパケット内におけるクラス識別子を解決し、個々のIPパケットフローを前記IP-QoSコードを割当てる割当手段を有することを特徴とする通信装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ISO参照モデルのネットワーク層のIPネットワークに用いられるマルチレイヤクラス識別通信装置と通信装置に関する。

【0002】

【従来の技術】

近年、インターネットはデファクトスタンダード的に国際的なネットワークとして標準化されており、TCP/IP (Transmission Control Protocol/Internet Protocol) がベースとして用いられている。通常、OSI参照モデルの7階層に当てはめれば、IPはネットワーク層として、TCPはトランスポート層としての働きを行い、最下位層のイーサネットやトーケンリングからIPに渡され、TCPからアプリケーション層に渡されてデータの取得等が行われる。

【0003】

ここで、ルータは、LAN間の中継装置として、リピータやブリッジとゲートウェイ間に配置され、通信回線から受け取ったフレームを蓄積し、フレームのデータ内にある宛先ネットワークアドレスに従って、適切な通信回線にフレームを転送する中継装置として用いられている。

【0004】

【発明が解決しようとする課題】

従来のルータは、基本的にパケット単位の優先制御を行わず、すべてのIPパケットを公平に扱っている。IPパケットは、ヘッダ部にIPアドレス等を含み、ルータのバッファに滞留するが、バッファからの読み出しは、FIFO (First In First Out) 方式による制御であるため、遅延優先制御はない。

【0005】

また、IPパケットが一定の出力ポートに集中すると、輻輳が発生し、ルータ内部のバッファで、パケット廃棄が発生する。通常、どのパケットを廃棄するかの廃棄優先制御は無い。

【0006】

翻って、周知のATM技術は、コネクションの概念（送信元、宛先を結ぶ経路をVPI/VCI (Virtual Path Identifier/Virtual Channel Identifier: 仮想パス・仮想チャネル識別子) により明示的に定義）を導入し、コネクション単位に必要なQoS（遅延特性、廃棄率等の通信品質: Quality of Service）を定義し、ネットワーク装置は各コネクションのQoSを満たすように優先制御を行う。

【0007】

(IP-QOS)

昨今、インターネット網上に簡易に優先制御を実現する技術が検討されている。代表的なものに、Intserv/RSVP (Resource Reservation Setup Protocol: ネットワークの帯域管理を実現するプロトコル) や、Diff-Serv (Differentiated Service: サービスの差別化) が挙げられる。Intserv/RSVPがATMのようなコネクションの概念を擬似的に導入するのに対し、Diff-Servは極力パケット単体の情報から優先制御を実現する技術である。

【0008】

Intserv/RSVPのような方式は、妥当なコストでは大規模基幹網に適応できない (Scalabilityに欠ける) という課題を有しており、普及していないのが現状である。これに対処するために考えられたDiff-Servは、(1) Scalableであること、(2) 妥当なコストで実現可能ということ、(3) OC-48クラス (Optical Carrier: 幹線用光ファイバのインターフェースで伝送速度別に) の高速化に対応できること、等が第一に考えられている。

【0009】

Diff-Servでは、ATM-QOSのような定量的なサービス保証をあきらめて、Best-Effort (最善の効果) からみて、相対的な品質差を設定し、簡易的にサービスの差別化を図る狙いがあり、IETF (Internet Engineering Task Force: インターネットの問題点を解決するための組織) の仕様化進行状況と、各ベンダー各社の対応状況をみても、非常に現実的な解として受け入れられている。

【0010】

(Diff-Serv)

Diff-Servは、インターネット技術の規定を行うIETFにて論議されている。Diff-Servは、サービスレベルの差別化を実現するための手段であり、QOS「保証」の手段ではない。あくまで（相対的な）優先制御の枠組みになる。つまり、Diff-Servは、QOSクラスのフレームワークを

定義するのみであり、各QOSクラスの詳細やクラス間のスケジューリング方式は、ベンダーマター、運用者マターとなっている。

【0011】

(Diff-Servにおけるサービスクラス)

Diff-Servでは、(1) Expedited Forwarding(Premium service)(EFクラス)、(2) Assured Forwarding Service(AF:前方保証クラス)、(3) Best Effort Service(BE:最大効果クラス)の3種類のトラヒック(サービスクラス)が規定されている。

【0012】

(1) Expedited Forwarding(EF:前方促進、プレミアム)クラスは、IP-CBR(Constant Bit Rate)のように、IP網上に仮想的な専用線を提供するものである。従って、UPC(Usage Parameter Control)やShaperも含んだ厳密な転送制御が必要となる。EFクラスは、帯域保証サービスとして位置づけられるため、後述するAssured Forwarding Service(AFクラス)や、Best Effort Service(BEクラス)に対して絶対的な優先度を持つ。

【0013】

(2) Assured Forwarding Service(AF)クラスは、Expedited Forwarding(Premium service)(EF)クラスと異なり、あくまで(相対的な)優先制御の枠組みになる。Assured Forwarding Serviceは、4種類の遅延クラスと、3種類の廃棄クラスを有する。遅延優先制御は、ネットワーク装置内におけるIPパケットの転送順序に優先度を持たせることで実現され、例えば、遅延に敏感なアプリケーションの転送遅延を下げる効果がある。廃棄制御はネットワーク装置内の輻輳発生箇所において、パケットの廃棄に優先度を持たせることで実現される。

【0014】

(3) Best Effort Service(BEクラス)は、EF、AFクラス以外のトラヒックが該当し、サービスにおける最も下位の優先制御を受ける。

【0015】

上記によって、Diff-Servのアウトラインを説明したが、Diff-Servの勧告は未だ流動的であり、上記の定義や運用方法が変更されることはある。

ありえる。

【0016】

そこで、本発明は、インターネットのIPパケットを扱うATM網のネットワーク層において、LAN間を接続するルータの機能に上記サービスクラスに応じて、パラメータの変更のみで柔軟に対応可能であり、汎用性に優れたマルチレイヤクラス識別通信装置を提供することを課題とする。

【0017】

【課題を解決するための手段】

本発明は、上記課題を解決するためになされたもので、IP (Internet Protocol) パケットを、ネットワークサービス品質 (Quality of Service: QoS) に応じて制御する手段を有した装置であり、下記機能を実現する手段を有することを特徴とする。

【0018】

(1) 入力インターフェースにおいて、受信したIPパケットのヘッダ情報、具体的にはIPヘッダに代表されるレイヤ3情報、およびその上位レイヤに相当するTCP/UDPヘッダに代表されるレイヤ4情報の組合せに基づいて、装置内におけるクラス識別子を解決し、個々のIPパケットフローを装置内で取り扱うサービス品質 (Internet Protocol Quality of Service: IP-QoS) コードを割当てる手段を有し、BEクラス、遅延の4種類と廃棄の3種類のAFクラス、EFクラスの計16種類にそれぞれ分けるために、CAM検索からSRA Mアクセスすることにより、優先度とスケジューリングに関するIP-QoSコードを割り当てる。

【0019】

(2) 上記IP-QoSコード割り当て (トラヒッククラスのmapping) については、汎用性をもたせるために、IPヘッダおよび、TCPヘッダ内容情報の任意の組み合わせで、後述する図5に示すように、優先度とフィルタリングによるクラス分けにより優先トラヒックを規定できる手段を有する。

【0020】

(3) 送信側のクラス別スケジューリング方式として、WRRS (Weighted

Round Robin Scheduling: 重み付け周回ロビンスケジューリング) と、固定優先スケジューリングを共存させ、各 I P - Q O S クラスを固定優先で選択できるとともに、最小帯域指定も可能とする手段を有する。

【0021】

(4) 受信側で解決した I P - Q O S コードは、パケット(主信号)データとは独立に、I P - Q O S クラス決定部で処理するので、装置内で取り扱う固定長のパケット情報格納セル(以下、便宜上“Object(オブジェクト)”と呼ぶこととする)に収容する手段を有する。

【0022】

(5) 宛先アドレスにスイッチするクロスバースイッチの入出力に備える入力インターフェース、および出力インターフェースが I P - Q O S クラス分けに応じて出力側でそのクラスに応じたスケジューリングを行うように、相互に連動して動作することにより、前記 Object 内に収容された I P - Q O S コードにより、所定の優先制御を行う手段を有する。

【0023】

(6) I P - Q O S コード単位にクラス分けされた I P パケットの群毎に、区別して処理することにより、同一クラスは同一処理が可能となって処理が迅速になっていくという統計多重効果を得るために、複数の I P フローを集約して共有できるクラスキュー管理を行う手段を有する。

【0024】

(7) 入力インターフェース、および出力インターフェースにおいて、I P - Q O S コード単位にトラヒックを監視し、I P - Q O S コード毎に廃棄処分可能か否かも判断でき、過剰なトラヒックの規制を行う手段を有する。

【0025】

また、本発明は、通信回線の複数の通信品質のレベルを選択可能な入力インターフェースと出力インターフェースとを備えた通信装置において、前記入力インターフェースは、入力する I P パケットから I P パケットヘッダ情報と T C P ヘッダ情報を抽出する I P パケット受信部と、前記 I P パケットヘッダ情報を検索キーとして I P - Q O S コード(クラス識別子) 解決メモリをアクセスして解決

対象となるクラス識別子を取得するIP-QOSクラス決定部と、前記IPパケット受信部より受信したIPパケットデータと前記IP-QOSクラス決定部において解決した該IPパケットに対応するIP-QOSクラスコードにより宛先特定の優先制御を行う受信側スイッチインターフェース制御部と、前記優先制御機能と送信要求を前記クラス識別子単位に前記出力インターフェースカードに送出する受信側スイッチインターフェースとを備え、前記IP-QOSクラス決定部は、前記IP-QOSクラス毎に設定された転送可能容量を超える過剰なトラヒックの流入を監視し、転送可能容量を超える場合は、該当するIPパケットを廃棄、あるいは、前記送信優先度を下げるポリシング（監視）制御を行う手段を有していることを特徴とする。

【0026】

また、本発明は、通信回線の複数の通信品質のレベルを選択可能な入力インターフェースと出力インターフェースとを備えた通信装置において、前記入力インターフェースから送出された優先制御されたIPパケットを受信する出力インターフェースは、受信した前記IPパケットを送信側ペイロードメモリに格納すると同時に該IPパケット情報を生成しFIFOメモリに書き込む送信側スイッチインターフェースと、前記IPパケット情報内に収容されたIP-QOSクラスコードに基づいて、クラス毎にキューイングを行い優先的に送信要求し且つWRR (Weighted Round Robin) 方式に基づいたスケジューリング機能により要求サービス品質を提供する制御機能を有するIP-QOSクラス別スケジューラと、前記送信側スイッチインターフェースからのIPパケットを前記IP-QOSクラス別スケジューラの優先度でスケジューリングされた順序に従って送出する送信側スイッチインターフェース制御部と、前記IPパケットを下位層のデータリンク層及びネットワークアクセス層に送出するIPパケット送信部とからなることを特徴とする。

【0027】

また、上記通信装置において、前記入力インターフェースに、受信した前記IPパケットのヘッダ情報(IPヘッダに代表されるレイヤ3情報、およびその上位レイヤに相当するTCP/UDPヘッダに代表されるレイヤ4情報の組合せ)

から前記IPパケット内におけるクラス識別子を解決し、個々のIPパケットフローを前記IP-QOSコードを割当てる割当手段を有することを特徴とする。

【0028】

【発明の実施の形態】

本発明の実施形態について、図面を参照して詳細に説明する。

【0029】

【第1の実施形態】

（1）構成の説明

図1は、本発明で適用するマルチレイヤクラス識別通信装置の構成を示すブロック図である。本実施形態による装置は、主として、入出力インターフェースカードと、N×Nクロスバースイッチ（1g）と、スイッチスケジューラ（1h）から構成される。

【0030】

入力インターフェースラインカードは、IPパケット受信部（1a）、受信側スイッチインターフェース制御部（1b）、受信側スイッチインターフェース（1d）、受信側ペイロードメモリ（1c）、IP-QOSクラス決定（1e）およびIP-QOSコード解決メモリ（1f）により構成される。

【0031】

出力インターフェースラインカードは、送信側スイッチインターフェース（1j）、送信側スイッチインターフェース制御部（1k）、IP-QOSクラス別スケジューラ（1m）、クラス別パケット情報キューイングメモリ（1p）、IPパケット送信部（1q）により構成される。

【0032】

なお、ここで、入力インターフェースラインカード及び出力インターフェースラインカードとしたのは、送信元アドレスから宛先アドレスに交換するクロスバースイッチ（1g）にカード形式に各必要な部材を装着したブロックを装着する意味であり、単にそれぞれ入力及び出力インターフェースとしてもよい。また、マルチレイヤクラス識別通信装置は、主にクロスバースイッチ（1g）に加えて入出力インターフェースとスイッチスケジューラ（1h）を備えたものである。また

、クロスバースイッチ（1 g）及びこのクロスバースイッチ（1 g）への切り換え制御するスイッチスケジューラ（1 h）を欠いた簡潔な構成の通信装置の場合には、インターネットでIPパケットを扱いネットワーク層までをサポートするルータや、物理層とデータリンク層の機能を備えパケットのフィルタリングなどを行うブリッジ等が、本実施形態の機能やプログラムソフトウェアを備えていてもよい。

【0033】

（入力インターフェースラインカード）

入力インターフェースラインカード（受信側）において、IPパケット受信部（1 a）は、OSI参照モデルのレイヤ4のトランスポート層のTCPによりアプリケーション層のパケットをIPパケットに分割して送出されたパケット情報から、IPパケットヘッダ情報と、およびその上位レイヤに収容されるTCP／UDPヘッダ情報を抽出し、その情報をIP-QoSクラス決定部（1 e）に転送する。

【0034】

ここで、IPパケット受信部（1 a）では、IPパケットから、受信したIPパケットの情報を元にCAMまたはSRAMエリアの内容によって定義されている各種条件を検索し、その条件にかなったアクションとしてそのIPパケットに対する、(i)Queueプライオリティ、(ii)廃棄/通過処理（フィルタリング）、(iii)SW優先/非優先処理、(iv)Diff servのDSCP値の付与など、の処理が導き出される仕組みを有している。

【0035】

IP-QoSクラス決定部（1 e）は、受信したIPパケットのヘッダ情報（IPヘッダに代表されるレイヤ3情報、およびその上位レイヤに相当するTCP／UDPヘッダに代表されるレイヤ4情報の組合せ）を検索キーとして、IP-QoSコード（クラス識別子）解決メモリ（1 f）をアクセスし、解決対象となるクラス識別子を取得し、受信側スイッチインターフェース制御部（1 b）にIP-QoSコードのクラス識別子を通知する。

【0036】

また、IP-QOSクラス決定部(1e)は、IP-QOSクラス毎に設定された転送可能容量を超えるような過剰なトラヒックの流入を監視し、転送可能容量を超える場合は、該当する(IP-QOSコードをもつ)パケットを廃棄、あるいは、転送優先度を下げる等のポリシング(監視)制御を行う手段を有しており、該パケットが属するIP-QOSコード毎に割り当てられた網リソース量に応じて転送パケットの送出頻度を決定する。

【0037】

受信側スイッチインターフェース制御部(1b)は、IPパケット受信部(1a)より受信したパケットデータと、IP-QOSクラス決定部(1e)において解決した該パケットに対応するIP-QOSクラスコードにより、クロスバースイッチ(1g)に対して優先制御を行う。

【0038】

受信側スイッチインターフェース(1d)は、送信要求をクラス単位および出力インターフェースカード単位に管理しており、常に高優先の送信要求を優先してスイッチスケジューラ(1h)に送り、この優先制御は、クロスバースイッチ(1g)に対する受信側ペイロードメモリ(1c)からの、パケット読み出しと対応し、遅延優先制御として位置づけ、IPパケットを宛先アドレスを選択・接続するクロスバースイッチ(1g)に出力する。

【0039】

受信側ペイロードデータメモリ(1c)は、入力されたIPパケット情報を格納し、送出する際に必要となる空きパケット情報を格納している。

【0040】

(出力インターフェースラインカード)

送信側スイッチインターフェース(1j)は、クロスバースイッチ(1g)から受信したパケットデータを送信側ペイロードメモリ(1i)に格納すると同時に、該パケット情報を生成し、FIFOメモリ(1n)に書き込む。左記パケット情報は、装置内でパケット単位に定義した仮想的な処理単位であり、以下、“オブジェクト”と呼ぶこととする。オブジェクトは、パケットデータを装置内で逐一持ち回ることを避けるために定義しており、パケットデータそのものではない

。装置内では、左記オブジェクトの受け渡しによって、パケット転送処理が実施されることとしている。

【0041】

FIFO (1n) は、送信側スイッチインターフェース (1j) からのパケットヘッダ情報を先読み先出しによって所定の遅延時間を保って送出する。

【0042】

IP-QoSクラス別スケジューラ (1m) は、オブジェクト内に収容されたIP-QoSクラスコードに基づいて、クラス毎にキューイングを行う。左記IP-QoSコードは、複数の遅延クラスおよび複数の廃棄クラスに対応づけされ、遅延優先順位の高いキューに格納されたオブジェクトが、優先的に送信側スイッチインターフェース制御部 (1k) に送出される。IP-QoSクラス別スケジューラ (1m) は、WRR (Weighted Round Robin) 方式に基づいたスケジューリング機能により、要求サービス品質が提供されるべく制御する機能を有している。また、IP網上で、仮想的な専用線を提供する特定クラス (Premium Service class) に対しては、UPC (Usage Parameter Control) やShaperも含んだ厳密な転送制御が実施される。

【0043】

送信側スイッチインターフェース制御部 (1k) は、送信側スイッチインターフェース (1j) からのIPパケットをIP-QoSクラス別スケジューラ (1m) の優先度等でスケジューリングされた順序に従ってIPパケット送信部 (1q) に出力する。

【0044】

IPパケット送信部 (1q) は、IPパケットを下位層のデータリンク層及びネットワークアクセス層のイーサネットやトーカンリングに出力する。

【0045】

(2) 動作の説明

図2は、本発明を適用する装置の機能を示した説明図である。図1で示した装置構成で、本発明で採用した優先制御の方法例について説明する。

【0046】

まず、IP-QOSクラス決定部（1e）においては、IP-QOSクラスの解決のため、レベルL3又はL4の所定情報をキーとして、CAM/SRAMをアクセスし、IP-QOSクラス（装置内のQOSコード）を解決する。この場合、BAクラス化/MFクラス化の両者をサポートする。

【0047】

また、EFクラスの追加と、ポリシング制御に関し、解決したIP-QOSクラス（EF/AF1～AF4/BE）に対するポリシング機能をサポートする。ポリシング制御の実態は、該クラス毎のトーカンと、送出パケット長の比較であり、トーカンがデータ長に達していない場合は、廃棄処理する。

【0048】

IP-QOSクラス決定部（1e）から受信側スイッチインターフェース制御部（1b）へ送出されるIP-QOSコード（クラス識別子）であり、IP-QOSコード解決メモリ（1f）を用いて、例えば、遅延、廃棄クラス種別として、（EF, H），（AF1～AF4, H/M/L），（BE, H/M/L）等により16種類（ $5 \times 3 + 1$ ）に振り分けられる。そのIP-QOSコード（クラス識別子）は、受信側スイッチインターフェース制御部（1b）と受信側スイッチインターフェース（1d）との間で、空きエリアのアドレスを取得する。

【0049】

また、IP-QOSクラス別スケジューラ部（1m）においては、IP-QOSクラス別スケジューラとして、受信側で解決したIP-QOSクラス（装置内のQOSコード）に基づくスケジューリングとして、実際は、高優先クラスのオブジェクトから、送信側スイッチインターフェース制御部（1k）に出力する処理の、EF>（AF1～AF4/BE）を実行する。ただし、（AF1～AF4/BE）は、WRRによるスケジューリングを実行する。

【0050】

また、IP-QOSクラス別スケジューラ部（1m）においての廃棄制御について、オブジェクトの引き落とし（drop）を生成するため、AF1～AF4/BEは、H/M/Lの3クラスのしきい値と、バッファ長の比較により、廃棄制御（Drop Objectの生成）を行い、EFについては、Hクラスの1種類を追加して

サポートする。

【0051】

また、E Fクラスの追加、シェイピング (Shaping) 機能に関し、解決した I P-QOSクラス (E Fのみ対象) に対するシェイピング (Delay Schedule) 機能が実行され、シェイピング制御の実態は、ポリシング制御と、同等のトーケンパケット方式であり、該クラス毎のトーケンと、送出パケット長の大小比較 (Compare (token length)) を行い、Tokenがlengthに達していない場合は、送出待ちとなる。

I P-QOSクラス別スケジューラ部 (1 m) はクラス別オブジェクトキューイングメモリSRAM (1 p) を用いて、オブジェクトに対するスケジューリングを行う。

【0052】

(優先制御)

装置内のスイッチ制御における優先制御として、以下の4形態を想定する。

【0053】

①受信側スイッチインターフェース制御部 (1 b) における装置内遅延クラスへのマッピング

装置内遅延クラスへのマッピングは、転送リクエストの送出方法で、2クラス (H/L) の遅延優先制御を指定→遅延クラス (6種類) に対応する。

【0054】

受信側スイッチインターフェース (1 d) はパケットを出力IFカード (ラインカード送信側) に転送する前に、スイッチスケジューラ (1 h) に接続要求 (request) を発行する。スイッチスケジューラ (1 h) は、各入力ラインカードからの転送要求を調停し、クロスバースイッチ (1 g) に対して、入出力バスの接続情報を通知する。また、各入力ラインカード (受信側) に対して、接続調停結果通知を発行する。

【0055】

受信側スイッチインターフェース (1 d) は、requestをクラス単位および出力IFカード単位に管理しており、常に高優先のrequestを優先して、スイッチス

ケジューラ (1 h) に送る。本優先制御は、クロスバースイッチ (1 g) に対する受信側ペイロードデータメモリ (1 c) からのパケット読み出しと対応し、遅延優先制御として位置づけられる。

【0056】

装置内遅延クラスへのマッピングは、受信側スイッチインターフェース (1 d) の前段に位置する受信側スイッチインターフェース制御 (1 b) がパケットのトラヒッククラスに応じて、本優先度を決定する。図2では、Diff-servに準拠したクラス数と整合をとるため、パケットのトラヒッククラスとして、6種類の遅延優先クラスを規定する。クロスバースイッチ (1 g) に対する2種類 (High/Low) のクラスキューと対応づける例を示している。

【0057】

②受信側ペイロードメモリメモリ (1 c) に入力パケットを書き込む際の廃棄クラスへのマッピング

この廃棄クラスへのマッピング空きエリアのアドレス (free page address) の取得処理において、3クラス (H/M/L) の廃棄優先制御を指定する。こうして、廃棄クラス (3種類) で、EFのみ (種類) に対応する。

【0058】

入力パケットを受信側ペイロードデータメモリ (1 c) に書き込む際に、前段の受信側スイッチインターフェース制御部 (1 b) は、受信側ペイロードデータメモリ (1 c) 内の空きエリアのアドレス (free page address) を取得し、左記アドレスにより、受信側ペイロードデータメモリ (1 c) に書き込む。

【0059】

空きエリアのアドレス取得に関して、受信側スイッチインターフェース (1 d) は、数種類の優先度を有している。説明を簡易にするため、図2中では、3クラス (高/中/低) = (H/M/L) の廃棄クラスを実装した例を示している。

【0060】

受信側ペイロードデータメモリ (1 c) の空きエリアのメモリ量を監視し、予め決められた閾値 (高) より少ない時は、高優先のパケット書き込みのみ許可する。閾値 (低) より少ない時は高優先、中優先のパケット書き込みのみ許可する

。それ以外の場合は、どのクラスに対しても書き込みを許可する。本制御は、メモリ(1c)に対する廃棄制御であり、トラヒッククラスの3種類の廃棄クラスと対応する。

【0061】

③送信側ペイロードメモリ(1i)から出力パケットを読み出す際の優先制御(遅延優先制御)

送信側ペイロードデータメモリ(1i)からのパケット読み出しは、IP-QoSクラス別スケジューラ(1m)、送信側スイッチインターフェース制御(1k)、送信側スイッチインターフェース(1j)により制御される。

【0062】

本装置の優先制御は、送信側ペイロードデータメモリ(1i)からのパケット読み出順序の制御であり、遅延優先制御に対応する。左記遅延優先制御は、IP-QoSコードにおける遅延クラス6種類と対応している。

【0063】

④送信側ペイロードメモリにスイッチから転送されてきたパケットを書き込む際の優先制御(廃棄優先制御)

クロスバースイッチ(1g)から転送されてくるパケットは、送信側ペイロードメモリに書き込まれる。書き込まれたパケットの諸情報は、オブジェクトとしてIP-QoSクラス別スケジューラ(1m)に通知される。IP-QoSクラス別スケジューラ(1m)は、クラス単位に送信側ペイロードデータメモリ内のキュー長を管理しており、キュー長と廃棄クラスの閾値比較により、送信側ペイロードメモリ内のパケットを廃棄するか否かの判定を行う。

【0064】

従って、本制御は廃棄優先制御を実行する。左記廃棄優先制御はトラヒッククラスの廃棄クラスと対応している。

【0065】

また、送信側スイッチインターフェース制御部(1k)では、2種類(送出/廃棄)のオブジェクトによるパケット送出/廃棄処理を実行し、送出の場合、FSUメモリに滞留しているパケットをリードコマンドを発行して読み出して送出し

、廃棄の場合、FSUメモリに滞留しているパケットをドロップコマンドを発行して廃棄する。

【0066】

(IP-QoSクラス決定部)

図3は、IP-QoSクラス決定部(1e)の要部説明図を示したものである。図3を参照してIP-QoSクラス決定部の詳細に関して説明する。

【0067】

同図に示す通り、IP-QoSクラス決定部(1e)は、ヘッダ抽出部(3a)、ヘッダチェック部(3b)、IP-QoSコード検索部(3c)、ポリシング制御部(3d)、IP-QoSコード出力部(3e)、パラメータレジスタ制御部(3f)から構成される。

【0068】

IPヘッダ抽出部(3a)において、図9に示すIPv4のIPヘッダとTCP/UDPヘッダのフォーマットから、所定の情報を抽出し、左記に含まれる各フィールドを(IPdataとして)IP-QoSコード検索部(3c)に送出する。

【0069】

なお、図9において、

(1) 細い斜線部(Ver)は、Check対象となるフィールドを示している

【0070】

(2) 粗い斜線部(TOS, Src IP Adress Dst, IP Adress, L4 Src Port, L4 Dst Port)は、クラスを特定するための(検索key対象となる)フィールドを示している。

【0071】

ここで、抽出される情報は、4ビット構成のバージョン(Ver)、8ビットで構成されるサービスタイプ識別子(TOS:Type of Service)、送信元IPアドレス(Src(Source) IP Address)、宛先IPアドレス(Dest(Destination) IP Address)、レイヤ4ヘッダ内のL4送信元ポート番号(L4 Src Port)、レイヤ

4ヘッダ内のL4宛先ポート番号 (L4 Dst Port) である。

【0072】

また、IPヘッダ自体の大きさを示すのIHL (Internet Header Length)、IPヘッダとIPデータを加えたパケット全体を示すTotal LengthであるDatagram Length、フラグメントを復元する際の識別子を示すIdentification、3ビット構成のフラグM、13ビットで構成され、分割されたフラグメントがオリジナルデータの何処に位置しているのかを示すFragment Offset、ネットワークに存在してもよい時間を示すTime to Live、上位層のプロトコルを特定するProtocol、IPヘッダのチェックサムを示すHeader Checksumをそれぞれ示している。

【0073】

IPヘッダチェック部 (3 b) は、IPヘッダの正常性を判定し、IP Header Error or Encap Fieldの結果を、IP-QOSコード出力部 (3 e) に出力する。IP-QOSコード検索部 (3 c) は、ヘッダ抽出部 (3 a) より受けた諸情報を検索キーとして、CAM (Contents Addressable Memory) およびSRAMをアクセスし、IP-QOSコードを解決する。この解決したIP-QOSコードには、装置内におけるクラス識別子、スイッチ制御に対する優先度等の情報が含まれる。

【0074】

ポリシング制御部 (3 d) は、前段のIP-QOSコード検索部において解決されたクラス種別単位にトラヒックを監視し、過剰なトラヒックの規制を行う。左記処理は、Token-Bucket方式によるトラヒック違反の監視である。Token-back et方式は、bucket内のtoken量は、経過時間T×平均レートrのT・rで増加する。パケット到着時に、該パケットに対して、パケットを入れる容器であるtokenに必要なtoken量に満たない場合は、該パケットを廃棄する。

【0075】

左記処理は、パケット長とtokenの比較であり、token < packet長の場合は、廃棄ビットを”ON”（廃棄対象パケットを明示）にして、後段ブロックに送出し、左記以外 (token \geq packet長) の場合は、廃棄ビットをスルーとする

簡易なロジックである。

【0076】

違反トラヒックの判定を高速化するために、ハードやソフトを作成・調整するインプリメンテーション上では、以下のようにしてもよい。

【0077】

入力パケットの通過条件を、「`token ≥ packet長`」ではなく、「`token >=0`」とする。

【0078】

パケットを送信した後、`token`からパケットサイズ分減算するため、`token`の値は負の値を取りうる。パケット送信時、`token`が負の値である場合は、送信パケットはトラヒック違反の対象とする。本判定回路を用いることにより、違反判定は、`token`の符号情報（1 bit）の情報ですむため、回路構成がシンプルになる。

【0079】

IP-QOSコード出力部（3e）は、IP-QOSコード検索部（3c）で解決したIP-QOSコードと共に、ポリシング制御部（3d）からのFilteringビット、および、ヘッダチェック部（3b）からのエラー情報をリタイミングして、後段ブロック（受信側スイッチインターフェース制御部：1b）に出力する。

トラヒッククラスのマッピング（MF／BA Classifier）IPヘッダに含まれる各フィールドをどのように組み合わせて、トラヒッククラスに対応づけるかは、運用者マターであり、各IETFのRFC等の勧告においても明確に記載されている訳ではない。この対応付けに、汎用性を持たせるためには、抽出したヘッダ情報の任意の組み合わせでトラヒッククラスにマッピングさせる必要がある。

【0080】

例えば、ある契約ユーザ間の特定のIPトラヒックに対して優先制御を行う場合は、送信元（Src）IPアドレスと、宛先（Dst）IPアドレスの組み合わせで、クラスを判断することになる。また、左記IPトラヒックの中でHTTP（WWWサーバで使うハイパーテキストをインターネット上で交換するプロト

コル) トラヒックのみ優先処理したいのであれば、さらにその上位レイヤのヘッダ内に収容されている Src Port #, Dest Port #を組み合わせてクラスを判断する必要がある。また、あるサーバから転送されるトラヒックのみに優先制御を行いたいのであれば、サーバのIPアドレスであるSrc IPアドレスと、Src Port番号のみを参照してクラスを判断する必要がある。このようにIPヘッダおよび、上位レイヤの複数のフィールドの組み合わせで、トラヒッククラスを判定する (Classifier) 方法を、Multi-Field (MF) Classifierと呼ぶ。

【0081】

上記に対して、Behavior Aggregate (BA) Classifierと呼ばれるクラス判定方法も存在する。BAクラス分け方法は、IPヘッダのTOSフィールドのみ参照して、トラヒッククラスを判定する方法である。IPヘッダに定義されるTOSフィールドは、Diff-Serv用に定義される特殊なフィールドである。TOSフィールドの目的は、各ルータでIPヘッダの各フィールドの組み合せから、トラヒッククラスを判定する処理を削減することにある。つまり、上流のルータがIPヘッダの各フィールドからトラヒッククラスを判定し、左記クラス情報をTOSフィールドに付与して転送する。次段以降のルータは、TOSフィールドのみを参照してトラヒッククラス別の優先制御を行えば良い。しかしながら、TOSフィールドの使用方法もベンダーマターになっているので、同一系列(同一ベンダー)の装置同士で接続するか、隣接するルータ間にて運用上の取り決めを行う等の制約が発生する。

【0082】

このように、トラヒッククラスのmapping (マッピング)においては、任意のフィールドの多種多様な組み合わせに応じて、登録できるような方式が必要であり、本発明では、以下に示す方法により、MF ClassifierとBA Classifierを同時にサポート可能である。

【0083】

(クラス検索部の動作)

クラス検索部の動作を図4、図5を参照して説明する。

図4は、クラス検索部（3c）の構成および動作を示したものであり、図5は検索処理のフローを示したものである。また、図4は、IP-QoSクラス検索部（3c）は、レイヤL4ポートの変換用SRAMと、部送信元IPアドレス検索用のCAMと、宛先IPアドレス検索用のCAMと、優先度マッピング検索用のCAMと、優先度を設定するためのSRAMとの各記憶部を備えたブロック図であり、ソフトウェア的に優先制御用のデータを抽出する。また、図5は、送信元アドレスと宛先アドレスとサービスタイプ識別子（TOS）とプロトコルとを入力し、検索したアドレスQを出力するCAM検索部と、検索したアドレスQと、上位層のTCP送信元ポートとTCP宛先ポートとを入力し、IP-QoSコードを出力するSRAMアクセス部とからなるフローチャートを示している。

【0084】

ここで、CAM上には主に条件情報、SRAMにはアクションの情報が記述される。SRAMに記述するアクションは、CAM上の条件記述と対応しており、特に媒体として分けることは、必要条件ではないが、SRAMの高速性を鑑みれば、CAMをSRAMで構成して、記憶領域を区分すればよい。あくまでも、CAMのリソース（エリア：エントリ数）を有効利用するためのインプリ的な手法を用いる。

【0085】

前述した通り、MF Classifierでは、送信元Src IPアドレス、宛先Dst IPアドレス、送信元Src Port番号、宛先Dst Port番号、プロトコル番号、TOSの任意の組み合わせから、トラヒッククラスを判定する。MF Classifierにおいて、最も簡易な検索方法は、entry毎にMaskを指定できるCAMに、登録情報として（送信元Src IPアドレス、宛先Dst IPアドレス、送信元Src Port番号、宛先Dst Port番号、プロトコル番号、TOS）の値を設定し、パケット入力毎にパケットヘッダ情報をもとに検索を行う方式である。しかしながら、CAMのbit幅などに制約があり、インプリ上の工夫が必要になる。

【0086】

そこで、本発明では、汎用的なCAMを使用し、且つ以下に示す2段階の検索

方法によりIP-QOSコードを取得するシーケンスを有する。なお、BA Classifierの処理(TOSフィールドのみの参照により、IP-QOSコードへ変換する処理)も、MFの処理と同じ枠組みの中で実施される。つまり、MFにおいて、TOSフィールドのみ有効となる運用を考えればよく、CAMの構造(Configuration)において、両者を共存させた運用が可能である。

【0087】

(1) まず、前処理として、個々のフィールドを“key値”に縮退し、CAM及びSRAMに登録する。

【0088】

(2) 次に、クラス判定処理として、前記“key値”により登録されたentryをCAM上で検索する。

【0089】

以下に、図4および図5に即して、具体的に説明する。

【0090】

(CAMへの事前登録)

前処理1：送信元(Src IP)アドレスの登録および縮退(図4中①)

予め、クラス判定処理のentryに含まれる送信元(Src IP)アドレス/prefixをすべて登録しておく。前処理において、特に送信元(Src IP)アドレス、宛先(Dest IP)アドレスを縮退する場合は、CIDR(Classless Inter Domain Routing)に基づく規定により、送信元(Src IP)アドレスをkeyとして、Longest Prefix Matchにより検索する。検索で返されたCAMアドレスをアドレスAとする。登録されていない場合は、“11”0”がアドレスA値となる。本処理は、図5のステップ(1)の処理となる。

【0091】

前処理2：宛先(Dest IP)アドレスの登録および縮退(図4中②)

前処理1と同様、予め、クラス判定処理のentryに含まれる宛先(Dest IP)アドレス/prefixをすべて登録しておく。宛先アドレスを検索keyとして、Longest Prefix Matchにより検索される。検索で返されたCAMアドレスをAddr_Bとする。登録されていない場合は、“11”0”がAddr_B

値となる。本処理は、図5のSTEP(2)の処理となる。

【0092】

なお、CAMをパラレル接続することにより、図5のSTEP(1)(2)は並列処理が可能である。

【0093】

前処理3：ポート番号（レイヤ4情報）の登録および縮退（図4中①'，②'）

本処理の目的は、Srcポート番号もしくはDestポート番号からレイヤ4アプリケーションを判定し、所定のkey値に縮退することである。ポート番号は、レイヤ4プロトコルを規定するWell Known Port番号と、各端末が任意に付与する番号に分類される。

【0094】

通常のレイヤ4アプリケーションにおいて、クライエントからサーバ方向に転送されるパケットのDest Port番号には、サーバで機能するレイヤ4アプリケーションのWell Known Port番号が付与されており、左記Well known番号をメモリ（テーブル）に登録する。

【0095】

運用で必要なWell Known Port番号は数種類に限られている（HTTP, TELNET, FTP等々）ため、メモリ量としては僅か（ $256 \times 8 = 2064$ bit）でよい。

【0096】

本処理では、ポート番号をアドレスとして、”ポート番号→Key”変換用のメモリにアクセスし、所定のkey値を得る。左記Keyは、レイヤ4を特定するcodeと、送信元(Src)ポートとして有効か、あるいは、宛先(Dest)ポートとして有効かを指定するflagより構成される。Src/Destポートのflagを指定する理由は、例えばサーバ向けのトラヒックは、高優先クラスとして扱うが、クライエント向けのトラヒックはBest Effort扱いする等の、片方向トラヒックに対するクラス分けも可能とさせるためである。

【0097】

図5のSTEP(3)およびSTEP(4)に示すように、本処理はSrcポート番号およびDestポート番号のそれぞれに対して実行される。Srcポート番号による変換キーの読み出し処理が、図5のSTEP(3)に相当し、Destポート番号による変換キーの読み出し処理が、図4における①'②'、および図5のSTEP(4)に相当する。

【0098】

両者の結果からWell Known Port番号が縮退された“Key値”、およびSrcポートに発生したか、Destポートに発生したかを示す“flag”が生成される。本処理が、図5のPory Keyを算出するSTEP(5)に相当する。

【0099】

(クラス判定処理(図4の③④))

前処理(図5におけるSTEP(1)～(5))で得られた各種Key値、およびTOS、プロトコル番号を元に、再度CAMを検索する。本処理が図5のStep(6)に該当し、各Keyの組み合わせ(どのKeyを有効情報として検索するか)は、CAM内のentry単位にmask情報として定義されている。本処理により、entryの格納されているCAMアドレス(addr_Q)が得られる。

【0100】

最後に、図5におけるSTEP(7)において、左記CAMアドレス(addr_Q)を外部メモリへのポインタ情報(アドレス)として使用して、アクセスすることにより、最終的に必要なトラヒッククラスの情報が得られる。

【0101】

図5に示す処理フローにて得られた情報(IP_INFO)は、受信側スイッチインタフェース制御(1b)部で、objectにmappingされ、及び受信側スイッチインタフェース(1d)に転送することにより、入力IFカードで実施される各種優先制御が行われる。

【0102】

なお、CAMの検索で登録にHitしない場合、CAMはアドレス“0”を返すこととし、この場合はBest Effortトラヒックを示すクラスコード

であると規定しておけばよい。外部メモリにおいても、リセット時の外部メモリ値は a l l ' 0 ' とし、アドレス" 0 " の領域には Best Effort トラヒック用のクラス情報が格納されることとなる。上記により、CAM検索で hit しないトラヒックは Best Effort トラヒックの扱いを受ける。各テーブルの詳細内容は、図6乃至図8に示す通りである。

【0103】

図6には、記憶手段のCAM (Contents Addressable Memory: 連想メモリ) に格納したテーブル構成例の表と、IP Src Prefix エントリの格納領域のテーブル図を示し、図7には、IP Des Prefix エントリ格納領域のテーブル図と、IPINFO 0アドレスエントリ格納領域のテーブル図を示している。また、図8はIPINFO エントリ格納領域のテーブル図である。

【0104】

なお、上記図6乃至図8の場合に、

D : Discard Indication (0:Normal, 1:Discard)

フィルタリング処理に使用する。

【0105】

P : Packet Priority (0: 低優先, 1: 高優先)

Route : 上位1ビット Route Change Flag

(0: Default Routeに設定、1: 下位4ビットのroute fieldに設定)

Output TOS

bit9 : DSCP Update Flag (0: TOSのDSCP Field (bit7-2) の書き換えを行わない、1: 書き換えを実施)

bit8 : CU Update Flag (0: TOSのCU Field (bit1, 0) の書き換えを行わない、1: 書き換えを実施)

bit7-2 : TOS DSCP (differentiated service code point) Field

bit1-0 : TOS CU (currently unused) Field

であることを付加する。

【0106】

(トラヒック規制手段)

本発明におけるトラヒック帰省手段は、図15に示すように、token bucket モデルによってトラヒック特性を表現し、(1)受信側インターフェース部におけるポリシング処理部(3d)、(2)送信側インターフェース部におけるシェーピング処理部(8f)において本token bucket モデルを適用する。

【0107】

つまり、トラヒック違反を検出するアルゴリズムは、シェーピング、ポリシング共に同等の回路構成となる。図15にToken bucket モデルを示す。Token bucket は、token 格納用レジスタ(バケツ:b)を持ち、周期的にバケツにtoken を追加する(平均レート(r)でtoken が増加する)。Token 加算処理においては、Token 上限値(b)と比較し、 $WC + token$ が b を越える場合は、b 値を設定する。

【0108】

パケット送信条件は、送信するパケットサイズ分のtoken がバケツに格納されていることである。また、パケットを送信した後、token からパケットサイズ分減算する。パケット送信時、パケットサイズ以上のtoken がbucket に無い場合は、該送信パケットはトラヒック違反の対象になる。ポリシング処理部(3d)に適用する場合は、違反の対象となったパケットを廃棄(あるいは低優先のマーキング等)を行い、シェーピング処理部(8f)に適用する場合は、違反したパケットを送信可能になる量のtoken がbucket に溜まるまで待ち合わせてから送信することとなる。

【0109】

(図16による補足説明)

トラヒック違反はパケットを転送するのに必要なtoken 量がbucket に溜まっていないときである。bucket 内token 量は経過時間T×平均レートrで増加する。パケット到着時に必要なtoken 量に満たない場合は、不足token が溜まるのに必要な時間を持ってからパケットを出力する。

【0110】

図16に、ポリシング・シェーピングの動作イメージを示す。同図に即して説

明すると、以下のようになる。

【0111】

(ポリシングの動作例)

(1) 時刻T1にパケット1が到着したとする。このとき、トラヒック違反でないので、同時刻T1'でパケットを出力する。出力した時点でのbucketに残っているtoken量はX1である。

【0112】

(2) 時刻T2にサイズs2のパケット2が到着したとする。パケット2を出力するのに必要なtoken量はs2であるが、 $s2 > X1 + (T2 - T1) \times r$ であったとする。この条件では、トラヒック違反であるので、パケット2を廃棄する。

【0113】

(シェーピングの動作例)

(1) 時刻T1にパケット1が到着したとする。このとき、トラヒック違反でないので、同時刻T1'でパケットを出力する。出力した時点でのbucketに残っているtoken量はX1である。

【0114】

(2) 時刻T2にサイズs2のパケット2が到着したとする。パケット2を出力するのに必要なtoken量はs2であるが、 $s2 > X1 + (T2 - T1) \times r$ であったとする。この条件では、トラヒック違反であるので、規定token量が溜まるまでの時間τだけ待ち合わせた後、パケット2を出力する。ここで、τは、 $s2 = X1 + \{(T2 + \tau) - T1\} \times r$ より算出可能である。

【0115】

(出力インターフェース(IF)カードにおける優先制御の詳細)

図2を参照しつつ出力インターフェース(IF)カードにおける優先制御について説明すれば、出力IFカードでは、IP-QoSクラス別スケジューラ(1m)が、object情報に定義されるIP-QoSコード(トラヒッククラス)を基に、所定の優先制御を行う。IP-QoSコードは、トラヒッククラスの識別情報として定義されるものであり、複数の遅延クラスと複数の廃棄クラスが表

現されている。IP-QOSクラス別スケジューラ(1m)で実施される優先制御は、図2で示した③、④の処理である。IP-QOSクラス別スケジューラ(1m)における優先制御の概念図を図10に示す。

【0116】

(IP-QOSクラス別スケジューラ(1m)における優先制御)

IP-QOSクラス別スケジューラ(1m)は、送信側スイッチインタフェース(1j)からFIFO(1n)経由でObjectを受け取り、Classifier(8a)にて、左記Objectを仕分けし、クラス単位のキューを振り分ける共有バッファ(8c)に格納する。左記共有バッファ(8c)は、クラス単位のObject格納用であり、それぞれ、EF、AF1~4、BEの計6種類のObject queue(8c)が存在し、共有バッファによる管理される。EF/AF/BEクラスの意味については、従来の技術にて説明したとおりである。

【0117】

EFクラスは、帯域保証サービスとして位置づけられるため、Shaper(8f)も含んだ厳密な転送制御を行う。図10中のWeighted Round Robin Controller(8d)で、EFクラスキューは、後述するAFクラスキューやBEクラスキューに対し絶対優先で読み出されるべく、Output Class Selector(8e)に対するセレクタ制御が実施される。Shaper(8f)は、Policing部と、同等のToken-bucketモデルによる回路構成であり、shaper部とポリシング制御部との差異は、違反パケットに対して、待ち合わせるか、廃棄するかの違いのみである。本箇所においても、インプリ上は、policing回路と同様、tokenに付与する符号bitのみの判定により、送出判定を高速化すると共に回路構成の簡素化を図ることも可能である。

【0118】

AFクラスおよびBEクラスに対しては、WRR(Weighted Round Robin)でスケジューリングされる。

【0119】

(WRRの補足説明)

WRRとは、Round Robin方式の拡張であり、予め決められたウェイト比にしたがって、各クラスに対してサービスを行う。WRRでは各クラス毎にカウンタを設けている。カウンタの値は、次にカウンタがリセットされるまでに送出可能なセル（or 固定長パケット）の個数を表し、リセット時には各クラスのウェイトの値に設定される。選択されたクラスのカウンタの値が0以上であり、且つバッファ内にセルが1つ以上あれば、そのクラスからセルを1つ送出し、カウンタ値をデクリメントする。もし、全てのクラスにおいて、ウェイトの値が0になるかバッファ内のセル数が0となった場合、全てのウェイトカウンタはリセットされる。従って全てのクラスに充分な入力トラヒックがある場合、出力されるセルの個数はウェイトの比に従う。

【0120】

AFクラスおよびBEクラスに対しては、左記クラス毎に、Weighted Counter (8g) を有し、Weight 値を保持している。Weighted Round Robin Controller (8d) は、左記Weight 値と共有バッファの各クラスキュー (8c) のキュー長の情報を元に、どのAFあるいはBEクラスのObject を読み出すか判定し、セレクタを制御する。選択されたAF及びBEクラスのキューは、FIFOの順序でObject を出力する。

【0121】

（廃棄制御）

図10中に示す廃棄制御ロジック部 (8b) は、図11で示されるロジックに基づき、廃棄制御が行われる。入力object 情報を元に、各クラスのキュー長 (8h) が管理される。各クラスには廃棄制御用のスレッシュホールド (8k) が定義されている。同時に、共有バッファ (8c) 全体のスレッシュホールドも定義されている。廃棄制御ロジック (8b) では、キュー長とスレッシュホールド (8k) の比較により廃棄制御を行い、廃棄するobject はDrop Objectとして送信側スイッチインターフェース (1j) に通知する。

【0122】

クラス単位のキュー長 (8h)、スレッシュホールド比較による廃棄制御に加え、共有メモリ (8c) 全体に渡る同一廃棄クラスのキュー長総和とスレッシュ

ホールド比較による廃棄制御も行われる。図11に示す通り、3種類の廃棄クラスに対してそれぞれ3種類のスレッシュホールドを定めている。従って、①全クラスで廃棄無し、②低優先クラスのみで廃棄、③低優先および中優先クラスで廃棄、④全クラスで廃棄の状態がある。図11において、*Thresh3*以上にキー長が増えると、空き領域があるのに関わらず、すべてのパケットが廃棄扱いとなる。

【0123】

同様に、AFクラス単位のみならず全メモリ内の廃棄クラスの総和に対しても、同様なスレッシュホールドを定義し、廃棄制御を行う。

【0124】

(WRR対象バッファ)

図14は、WRR対象バッファの要部説明図である。各クラス別objectキー(8c)に対して、現状のウェイト値を管理するカウンタ(111, 121, …1n1)と共に、ウェイトカウンタリセット処理後のウェイト値を管理する予備判定用のウェイトカウンタ(112, 122, …1n2)を設けている。

【0125】

最初の判定におけるWeighted Round Robinで即時に判定できない時は、ウェイトカウンタリセット処理後のウェイト値(予備判定用のウェイトカウンタ)で同様にクラス判定を行う。左記処理でも判定できない場合は、固定遅延優先でAFクラスを選択する。左記クラス判定処理は、ループ処理を削除した並列処理で実施されるため、高速性を損なうことはない。

【0126】

(スケジューリング処理フロー)

図12及び図13に本スケジューリングの処理フローを示す。処理のパスとしては、EFクラス判定パスと、AF/BEクラス判定のパスが存在する。EFクラスがAF/BEより絶対優先であるので、EFクラスの送出要求があるときは、EFクラスが選択され、EFクラスの送出要求がないときのみ、AF/BEの所定クラスが採用される。

【0127】

AF/BEクラス判定に関しては、2度目までの判定まではWeighted Round Robinで実現し、3度目の判定として固定遅延優先で選択される。

【0128】

なお、図12及び図13の処理フローを読み込む場合は、以下の点を把握してほしい。図12及び図13において、各パラメータを説明すれば、EFクラスでは、W[EF] : EFクラスの初期ウェイト（加算token値）、WH[EF] : tokenの上限値、WC[EF] : EFクラスのtoken値（変数）、AddTime[EF] : token加算周期（例えば、契約時の平均レートで規定）、であり、AF/BEクラス(i AF1, 2, 3, 4, BE)では、W[i] : クラス毎初期ウェイト値、WH[i] : ウェイトカウンタ上限値、WC[i] : ウェイトカウンタ値（変数）、WC[i]r : 1回のウェイトカウンタリセット後のウェイトカウンタ値（変数）、WC[i]r \leq WC[i] + class ptr : 送出、検索クラスを指すポインタ、MPSZ (Maximum Packet Size) : AF/BEクラスで適用：レアケースの判断基準とするパケット最大長。fixed Priority(Q[i]) : WC値に依存せず、Q[i] > 0 : キュー内のobjectの有無のみをみて固定優先度（AF1, 2, 3, 4, BEの順）に基づいた出力要求を送信する。EF/AF/BE共通として、Length : パケット長である。

【0129】

まず、EFクラス選択の処理過程を説明する。

【0130】

(Tokenの周期的な加算処理)

図12において、初期化処理について、ステップS0aにて、token(WC[EF])を“0”に初期化する（ステップS0a）。ステップS0bにて、Token加算タイミングを生成し、所定タイミング(cnt == AddTime[EF])であれば、S0cにてTokenを加算する。現状のWC[EF]に初期tokenであるW[EF]を加算する。つぎに、Token加算周期は、例えば契約時における平均レートで規定される。上記処理において、Token上限値(WH[EF])と比較し、WC[EF] + W[EF]がWH[EF]を越える場合は、WH[EF]を設定する。

【0131】

(E F クラス送信要求の生成)

次に、図13に示すように、ステップS1aにて、Object読み出し要求を受信した場合は、S1bにてキュー長の判定を行い、S1cにて、Tokenの有無判定を行う。

【0132】

左記判定において、EFクラスのキューにObjectが1個以上あり、且つtokenも存在する場合(0以上)には、S1dにて、出力要求(Send_req [EF] = ON)を送出する。その後S1eにてパケット送出後、tokenからパケット長分減算する。S3の最終段のセレクタ処理にて、最優先クラスであるEFクラスの送出要求が受理される。

【0133】

次にAF/BEクラス選択処理について説明する。

【0134】

(AF/BEクラス選択処理)

このAF/BEクラス選択処理の場合、図12と図13とを参照して説明する。

【0135】

(1) S0d、およびS0eにて、クラスポインタ、およびウェイトカウンタの値を初期化する。

【0136】

(2) S1aにて、Object読み出し要求を受信し、EFクラスが送出不可のとき(S1b、S1dがいずれもNoのとき)のみ、AF/BEのクラス選択処理が実施される。S2bにてキュー長の判定を行い、S1cにて、weightの有効性を判定する。

【0137】

(3) AF/BEクラスは複数クラスあるため、S2eにて、全クラスに対して、上記(2)の判定処理を行う。S2gにて、クラスポインタ(検索クラスを指すポインタ)の更新は、ラウンドロビンの規則に従う。

【0138】

(4) 送出条件（キュー長>0 且つ Weight 値>0）に合致するクラスが存在する場合は、S2d にて該当クラスの出力要求（Send_req [Class] = ON）を送出する。

【0139】

前述した（1）～（4）の処理が1st ルーチンによるクラス選択である。

【0140】

(5) 前述した（3）の検索処理において、全てのクラスについて、送出条件に合致しない場合は、全クラスに対して、クラス毎に規定された初期ウェイト値を現状のウェイト値に加算する。（左記処理をウェイトカウンタリセット処理と呼んでいる。）

(6) 1st ルーチンでクラス選択ができない（出力クラス候補が存在しない）場合は、（5）において1回のウェイトカウンタリセット後のウェイト値によって再度クラス選択を行う。図11に示すように、方法として、現状のウェイト値を管理するとレジスタ（WC [class]）と共に、1回のウェイトカウンタリセット処理後のウェイト値を管理するレジスタ（WC [class] r）を別に有し、（4）の処理と同様、送出条件（キュー長>0 且つ WC [class] r >0）に合致するクラスが存在する場合は、該当クラスの出力要求（Send_req [Class] r = ON）を送出する。

【0141】

前述した（5）～（6）の処理が、2nd ルーチンによるクラス選択である。

【0142】

(7) 2nd ルーチンにおいても、クラス選択ができない（出力クラス候補が存在しない）場合は、特例措置として、固定優先的にクラス選択を実施する。本処理の意図は、回線容量に空きがあるにも関わらず、パケットが送出できない状態に陥る状態を救済するためであり、予め決められている固定優先度（例えば、図8のモデルでは、AF1>AF2>AF3>AF4>BEの順）に基づいて、送出クラスを選択する。左記処理は、ウェイト値に依存せず、キュー内のobject の有無のみを判定して出力候補を決定する手法であり、本手法を用いる

ことにより、1回のクラス選択処理においてウェイトカウンタリセット処理が2度以上起こるケースはない。通常の判定は、Weighted Round Robinで実現し、特例の判定として固定遅延優先（図12中では、Fixed Priority MODEと呼ぶ。）で実現する。

【0143】

（8）ステップS2nにて、（4），（6），（7）で処理された該当クラスの出力要求より、出力クラスを選択する。

【0144】

本ステップの処理において、パラメータとして最大パケット長（Maximum Packet Size: MPSZ）を導入する。最大パケット長導入の意図は、過大なパケット長が到着したときにウェイトカウンタの値が負の方向に大きな値となることを救済するための措置である。例えば、IPパケットの最大長は、64kbyteであるが、左記の様なパケットは、異常な事態として考えて良い。統計的に、一般的なパケット長は数kbyteまでを想定すれば十分である。そこで、異常パケットとみなされる最大パケット長を設定可能とし、過大なパケットが到着したことを、パケット長（Length）と最大パケット長との比較を行う。パケット長>最大パケット長の場合は、特殊ケースとして扱い、ウェイトカウンタの減算処理をスキップする、もしくは、ある一定のペナルティを与えるべくウェイト値を強制的に”0”にする等の処置を行う。

【0145】

上述した実施形態によって、未だ規格的に流動的なDiff-Servにおいても、OSI参照モデルのTCP層とIP層の組み合わせによって、IPパケットから、IPパケットとは独立して、IP-QoSコードをそれぞれ割り当て、IP-QoSコードに応じて各種の通信品質のクラス分けを行うことが望まれる。このIP-QoSコードのクラス分けによって、通信系の輻輳状態を解消することが大きな効果として呈示し得る。

【0146】

【発明の効果】

本発明による効果は、以下の点にある。

【0147】

(1) オペレータの設定によりパケットのレイヤ3・レイヤ4情報の組み合わせ（上位アプリケーション別フロー単位）による優先トラヒックを規定できる。

【0148】

(2) 汎用的な運用を想定して、WRRS (Weighted Round Robin Scheduling) と固定優先スケジューリングとを組み合わせることが可能である。各QoSクラスを固定優先で選択できるとともに、最小帯域指定も可能である。

【図面の簡単な説明】

【図1】

本発明によるマルチレイヤクラス識別通信装置の構成図である。

【図2】

本発明によるマルチレイヤクラス識別通信装置の構成動作の説明図である。

【図3】

本発明によるIP-QoSクラス決定部の要部説明図である。

【図4】

本発明によるクラス検索部の動作説明図である。

【図5】

本発明によるクラス検索部の処理フローチャートである。

【図6】

本発明によるCAM領域分割のテーブル構成例及びIP Src Prefixエントリ格納領域の構造図である。

【図7】

本発明によるIP Src Prefixエントリ格納領域のテーブル構成図及びIP Src Prefixエントリ格納領域の構成図である。

【図8】

本発明によるIPINFOエントリ格納領域のテーブル構成図である。

【図9】

本発明に用いるIPv4 & TCP/UDP/Other Header フォーマット図である。

【図10】

本発明によるIP-QoSクラス別スケジューラの原理図である。

【図11】

本発明による廃棄制御ロジックの説明図である。

【図12】

本発明によるスケジューラの処理フローチャートである。

【図13】

本発明によるスケジューラの処理フローチャートである。

【図14】

本発明によるWR R対象バッファ要部の説明図である。

【図15】

本発明によるtoken-bucketモデルによるポリシング・シェーピング図である。

【図16】

本発明によるポリシング・シェーピング動作の説明図である。

【符号の説明】

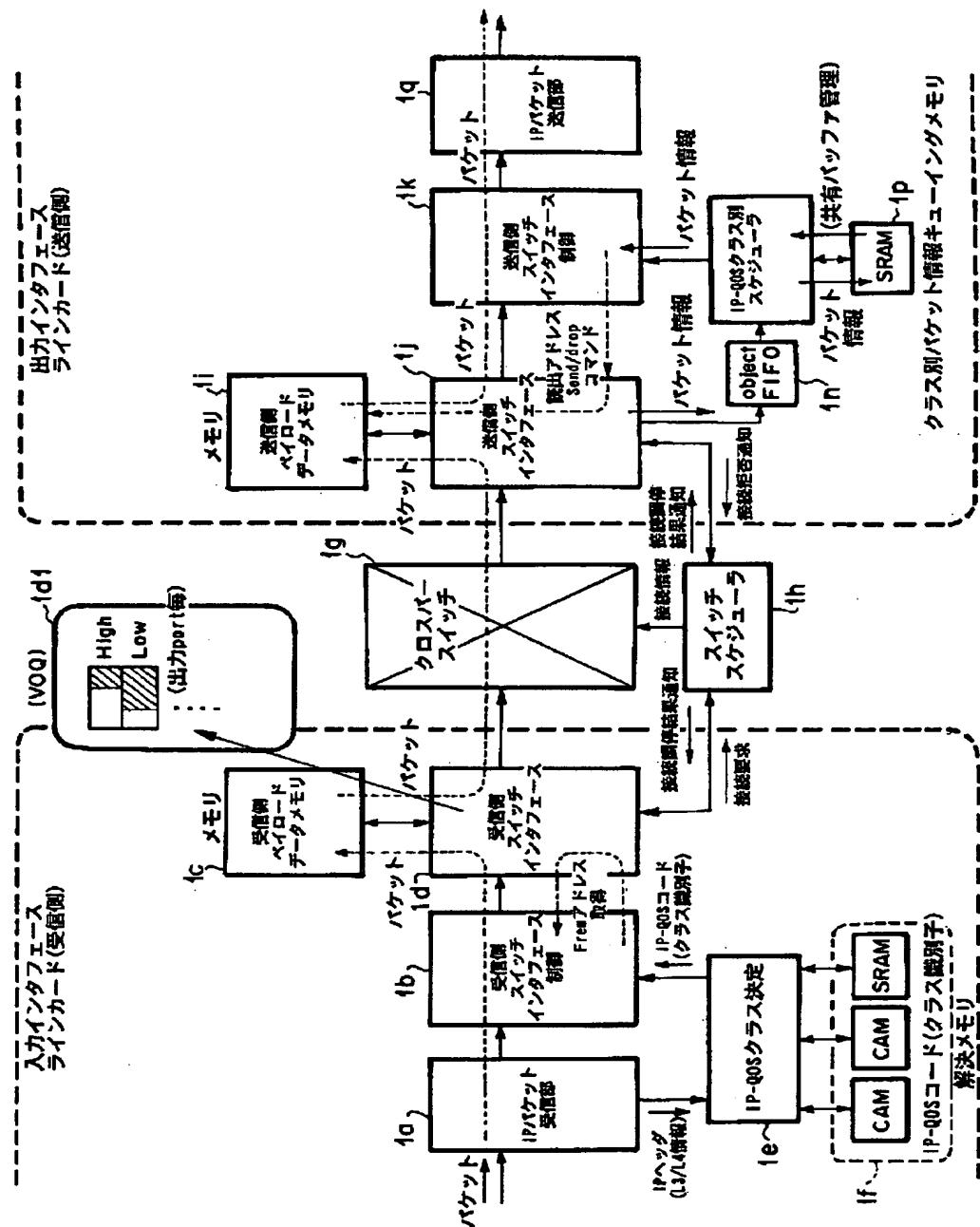
- 1 a IPパケット受信部
- 1 b 受信側スイッチインタフェース制御部
- 1 c 受信側ペイロードデータメモリ
- 1 d 受信側スイッチインタフェース部
- 1 e IP-QoSクラス決定部
- 1 f IP-QoSコード(クラス識別子)解決メモリ
- 1 g クロスバースイッチ
- 1 h スイッチスケジューラ
- 1 i 送信側ペイロードデータメモリ
- 1 j 送信側スイッチインタフェース部
- 1 k 送信側スイッチインタフェース制御部
- 1 m IP-QoSクラス別スケジューラ部
- 1 n FIFO
- 1 p SRAM

- 1 q IPパケット送信部
- 3 a ヘッダ抽出部
- 3 b ヘッダチェック部
- 3 c IP-QOSコード検索部
- 3 d ポリシング制御部
- 3 e IP-QOSコード出力部
- 3 f パラメータレジスタ制御部
- 8 a クラスファイア
- 8 b 廃棄制御ロジック
- 8 c 共有バッファ
- 8 d 重み付けラウンド・ロビン制御部
- 8 e 出力クラス選択部
- 8 f シェイパー
- 8 g 重み付けカウンタ
- 8 h クラスキューの長さ
- 8 j 全キュー長さ
- 8 k スレシホールド(クラス、レベル、全)

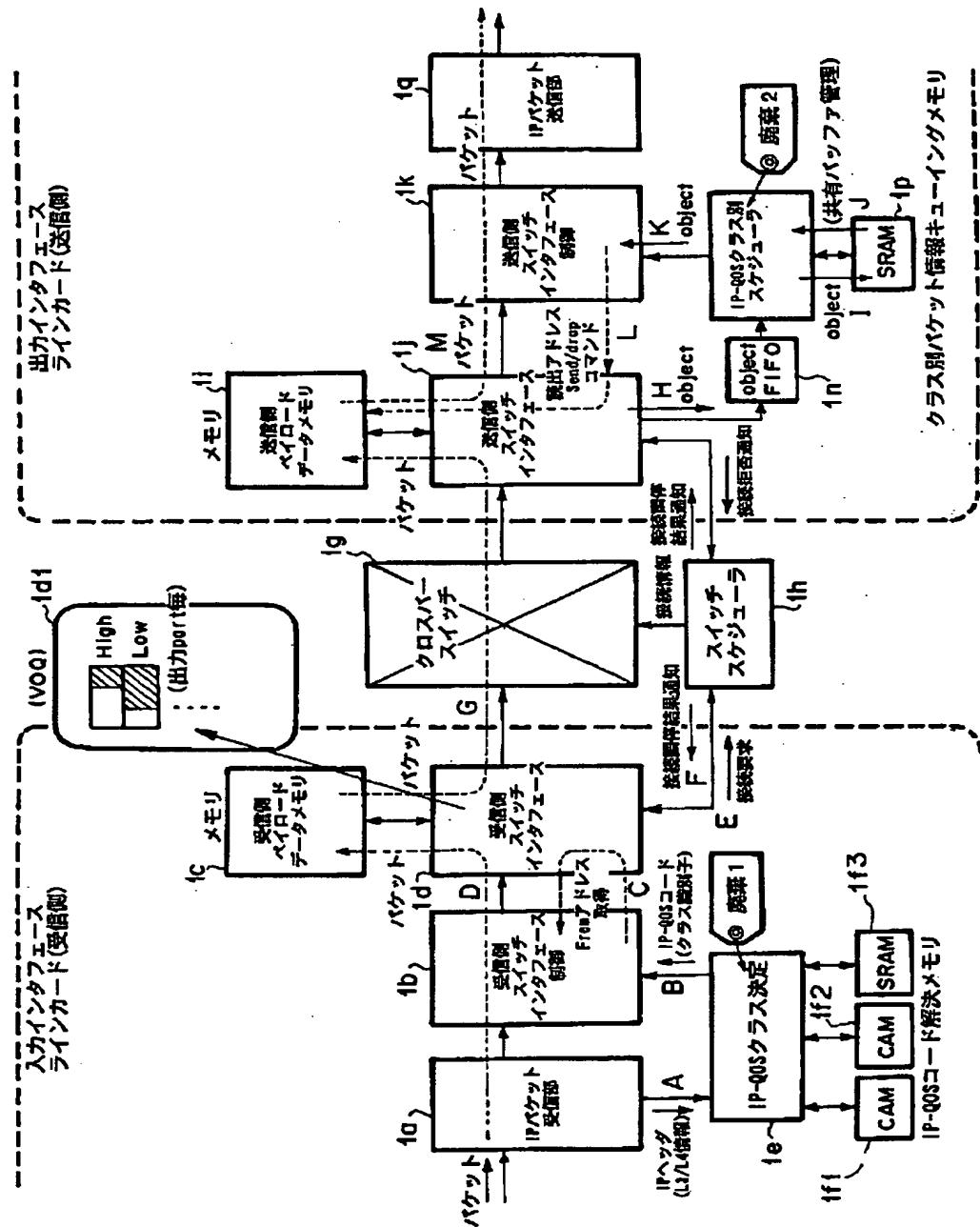
【書類名】

図面

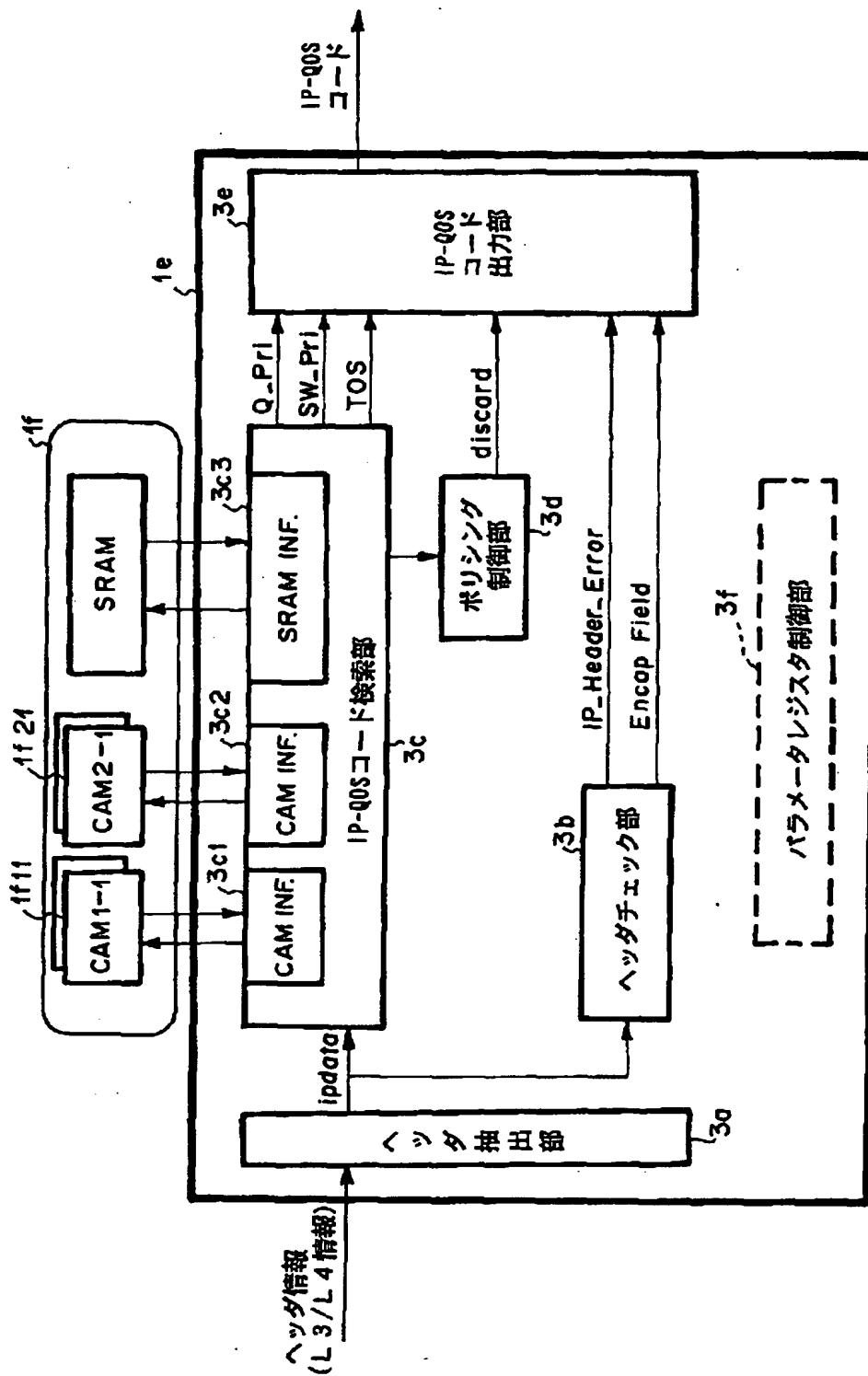
【図1】



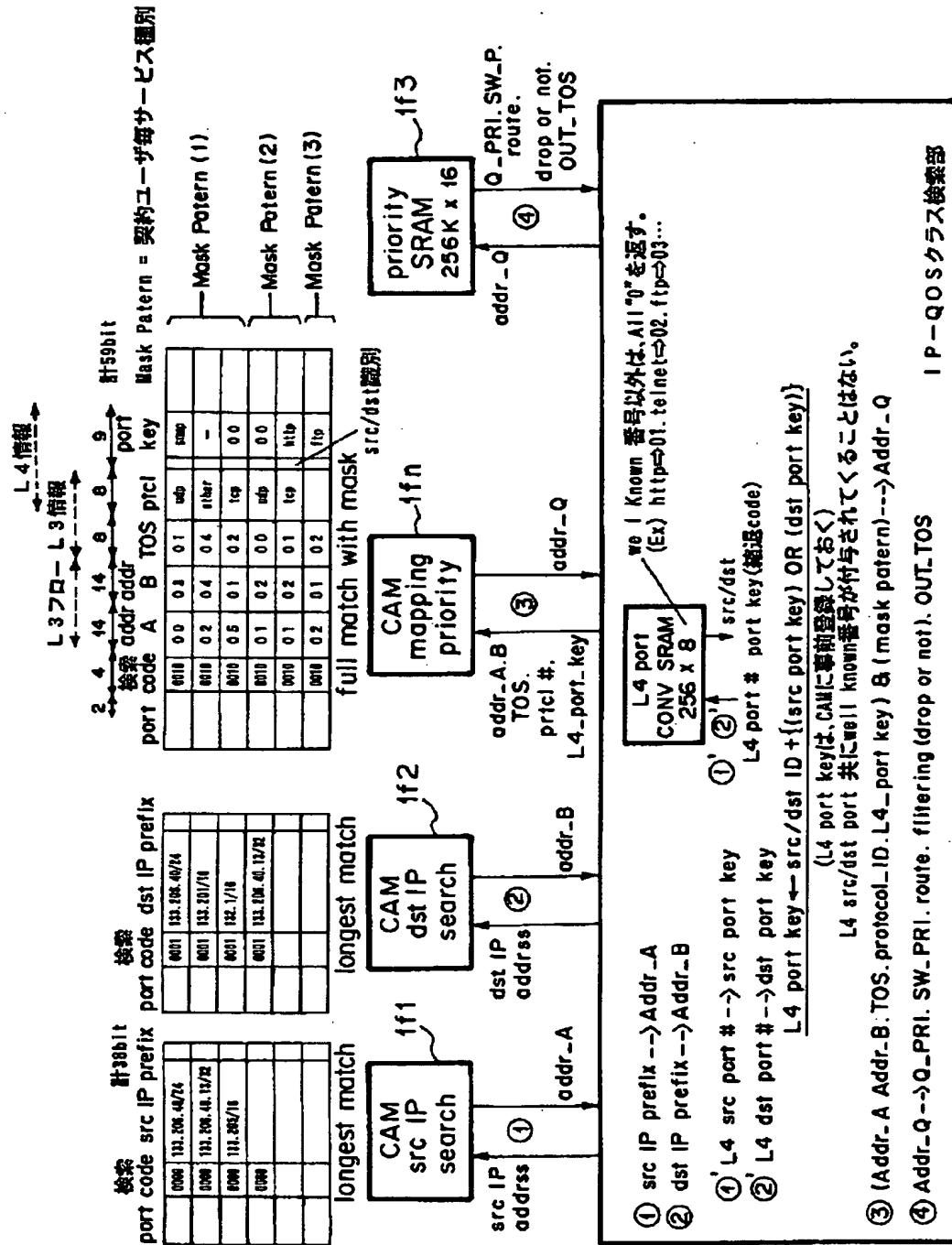
【図2】



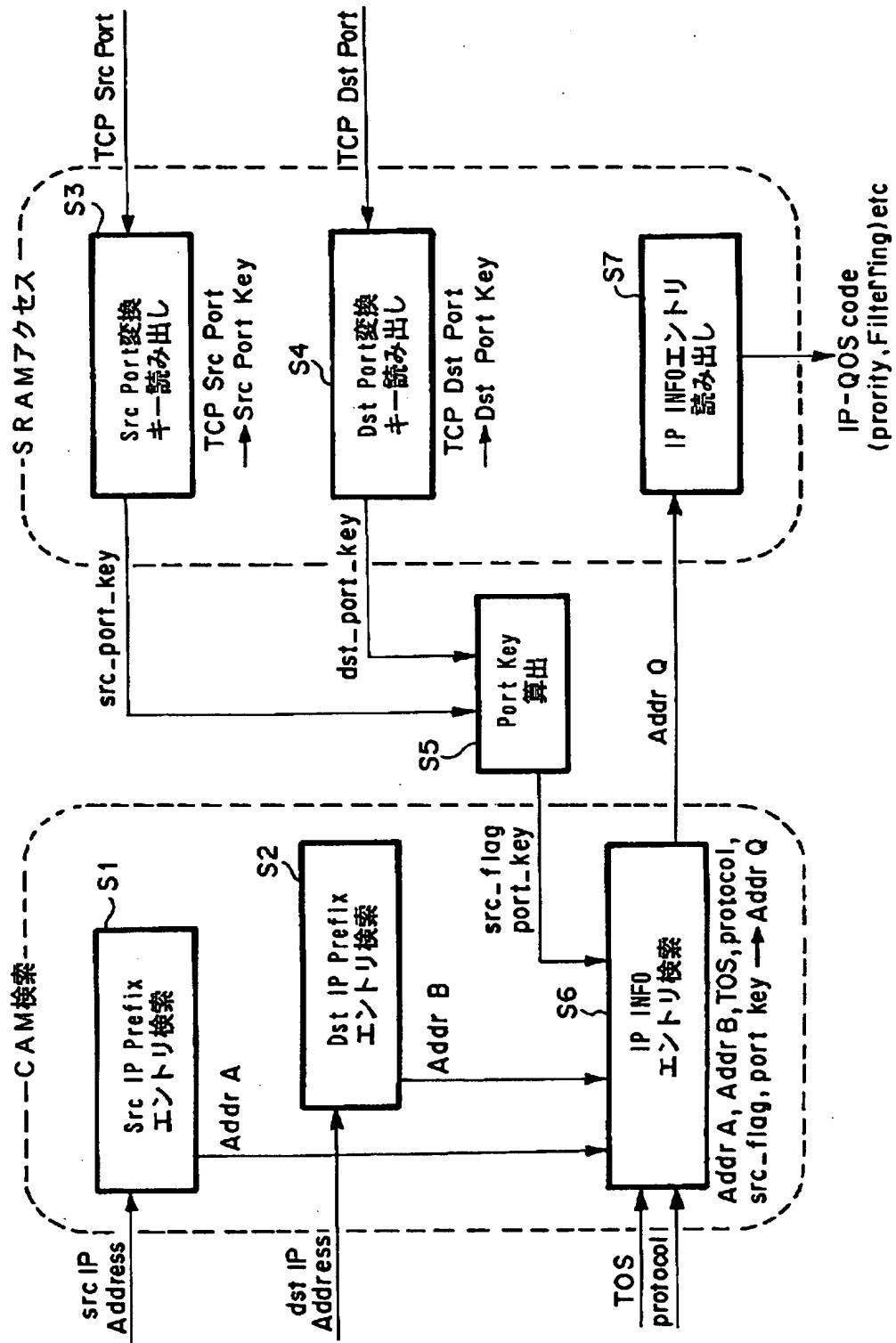
【図3】



[図4]



[図5]



【図6】

【CAM 領域分割】

CAMアドレス	CAMデータ (最大64bit)	マスクパターン(64)	サーチ方法
Addr_A～	IP Src Prefixエントリ格納領域		Longest match
Addr_B～	IP Dst Prefixエントリ格納領域		Longest match
Addr_Q～	IPINFO検索エントリ 格納領域		Full match With MASK

【I. IP Src Prefixエントリ格納領域 : 検索Code=0000】

CAMアドレス (Addr_A)	CAMデータ (38bit)			
	HW #(2)	検索 Code(4)	IP Src Address /Prefix (3 2)	未使用(26)
A #1	00	0000	IP Src Address #1 / Prefix	
A #2	00	0000	IP Src Address #2 / Prefix	
A #3	01	0000	IP Src Address #1 / Prefix	
:	:	:	:	

【図7】

【2. IP Dst Prefixエントリ格納領域 : 検索Code=0001】

CAMアドレス (Addr_B)	CAMデータ (38bit)			
	HW #(2)	検索 Code(4)	IP Dst Address /Prefix (32)	未使用(26)
B #1	00	0001	IP Dst Address #1 / Prefix	
B #2	00	0001	IP Dst Address #2 / Prefix	
B #3	01	0001	IP Dst Address #1 / Prefix	
:	:	:	:	

【3. IPINFOアドレスエントリ 格納領域 : 検索Code=0010】

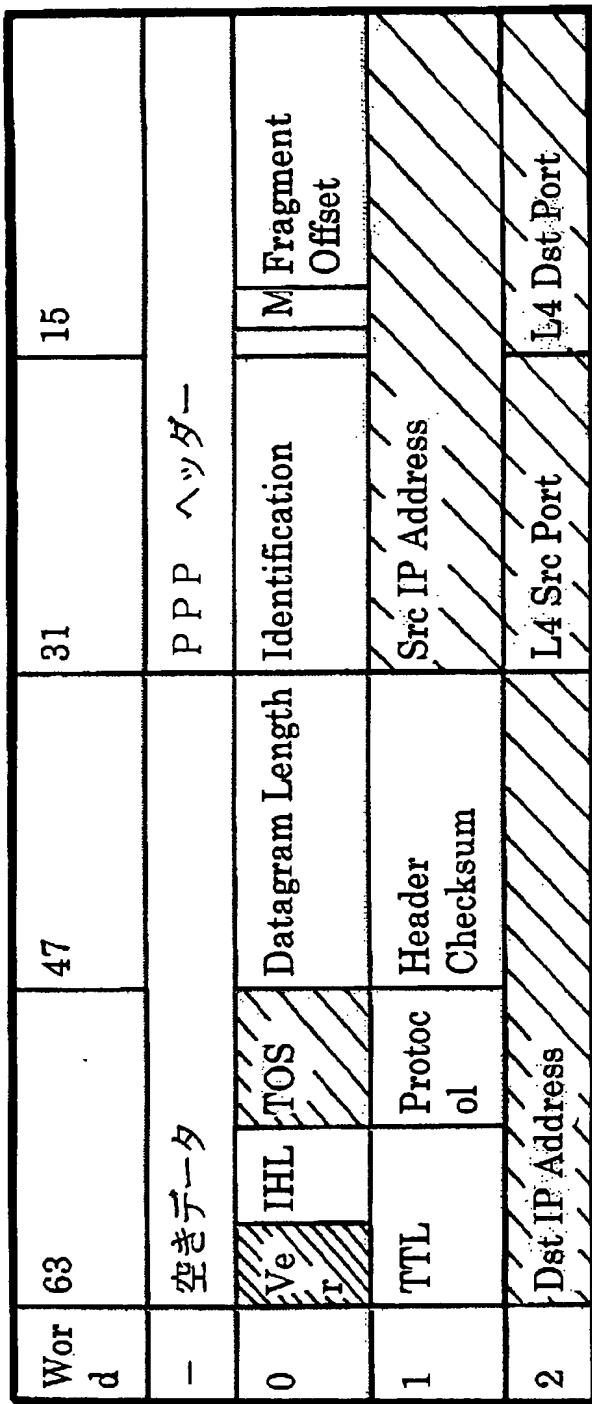
CAMアドレス (Addr_Q)	CAMデータ (55bit)									
	HW #(2)	検索 Code(4)	Addr A (14)	Addr B (14)	ToS (8)	Proto col # (8)	Src/D st (1)	Port key(8)	未 (5)	
Q #1	00	0010	A1	B1	01	Tcp	S	http		
Q #2	00	0010	A1	B2	04	Udp	D	Snmp		
Q #3	01	0010	A3	B1	02	Tcp	S	ftp		
:	:	:	:	:	:	:	:	:		

【IPINFOエンタリ】

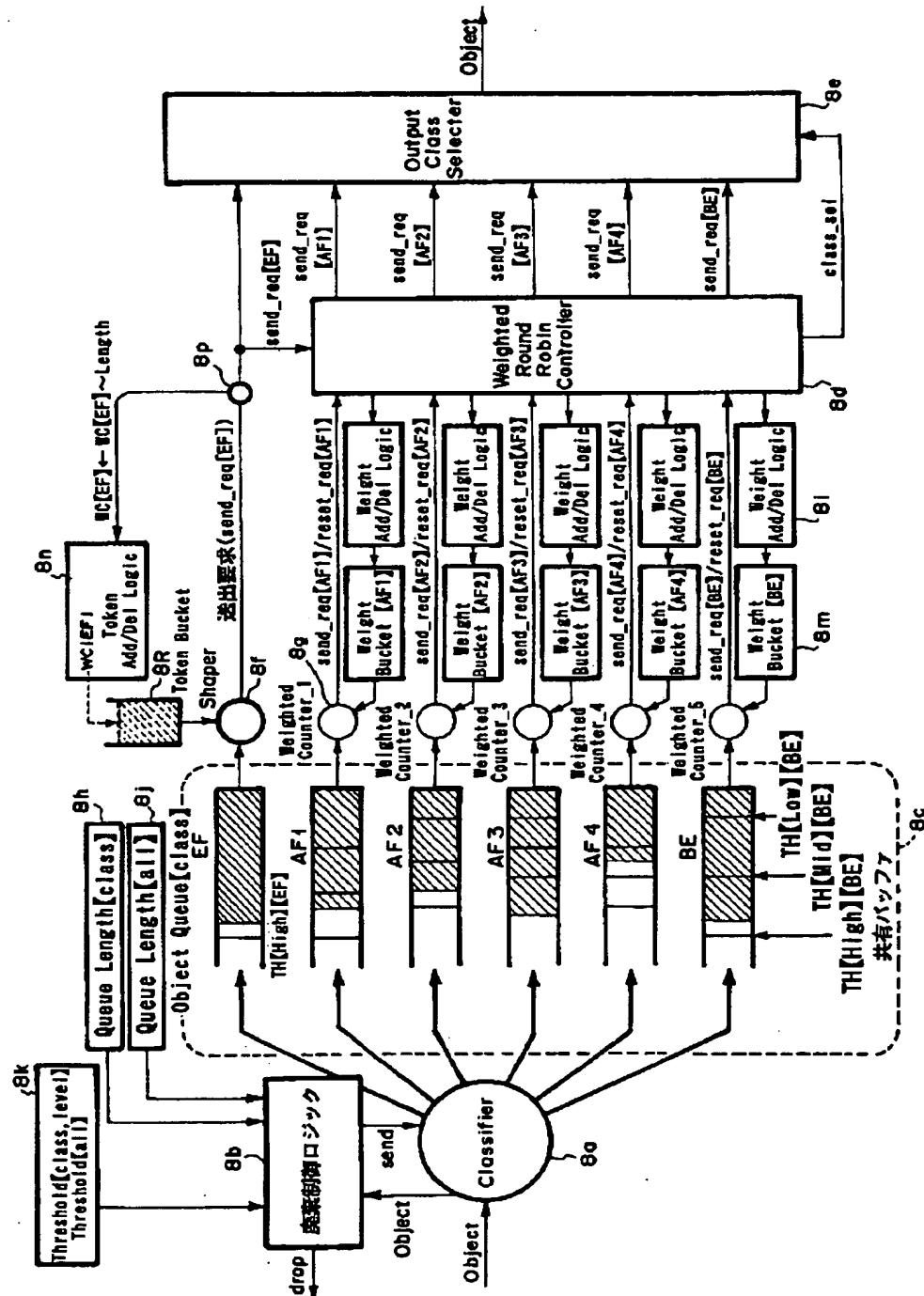
アドレス(16bit) : 上位2bit =00		データ(24bit)					
下位14 bit = Hit Addr_Q		Q_pri(4)	D	P	Route(1+4)	Output ToS(2+8)	Reserve (3)
Addr Q0		0000	0	0	0000	1 1	011011 00
Addr Q1		1101	0	1	0	0000	1 1
Addr Q2		1101	0	0	0	0000	0 0
:		:	:	:	:	:	
Addr Qi		1110	0	1	1	0101	0 0
:		:	:	:	:	:	

【図9】

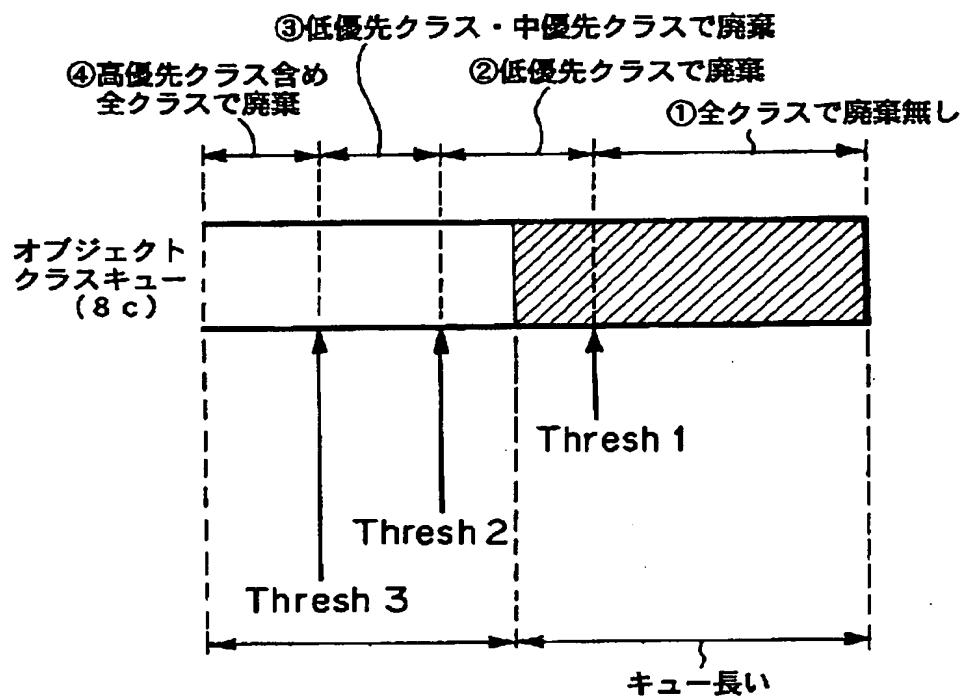
(IPv4 & TCP/UDP/Other Header フォーマット)



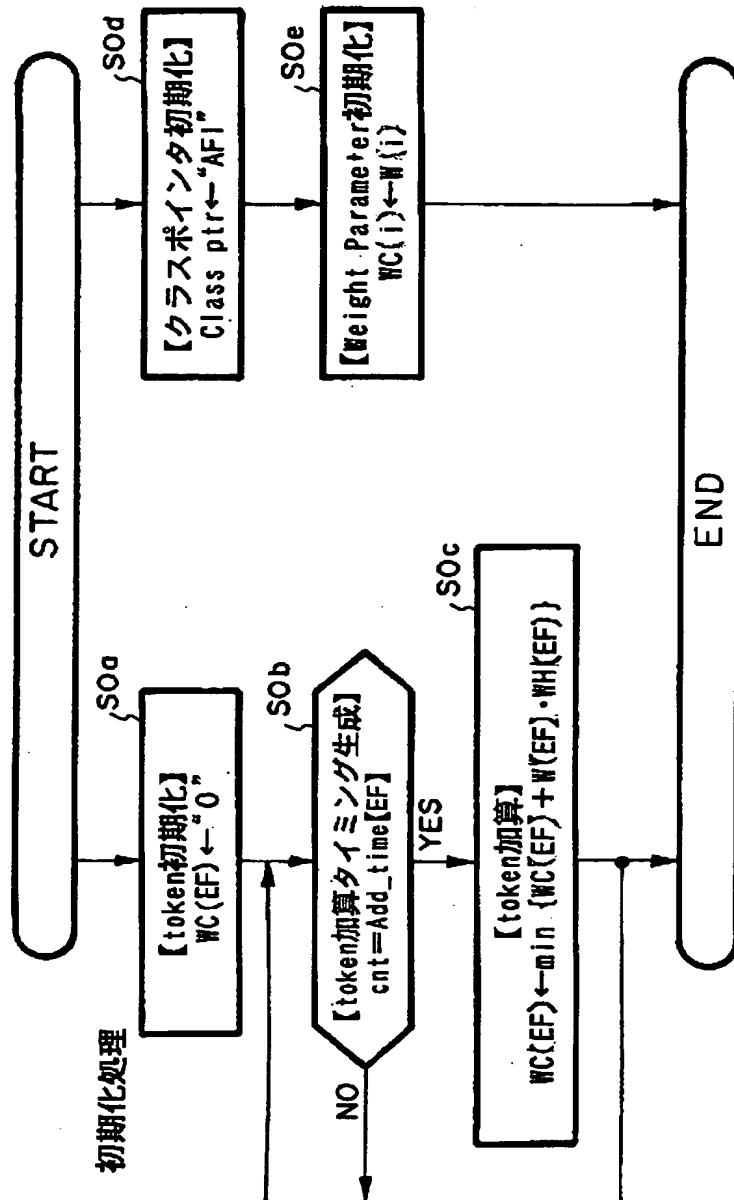
【図10】



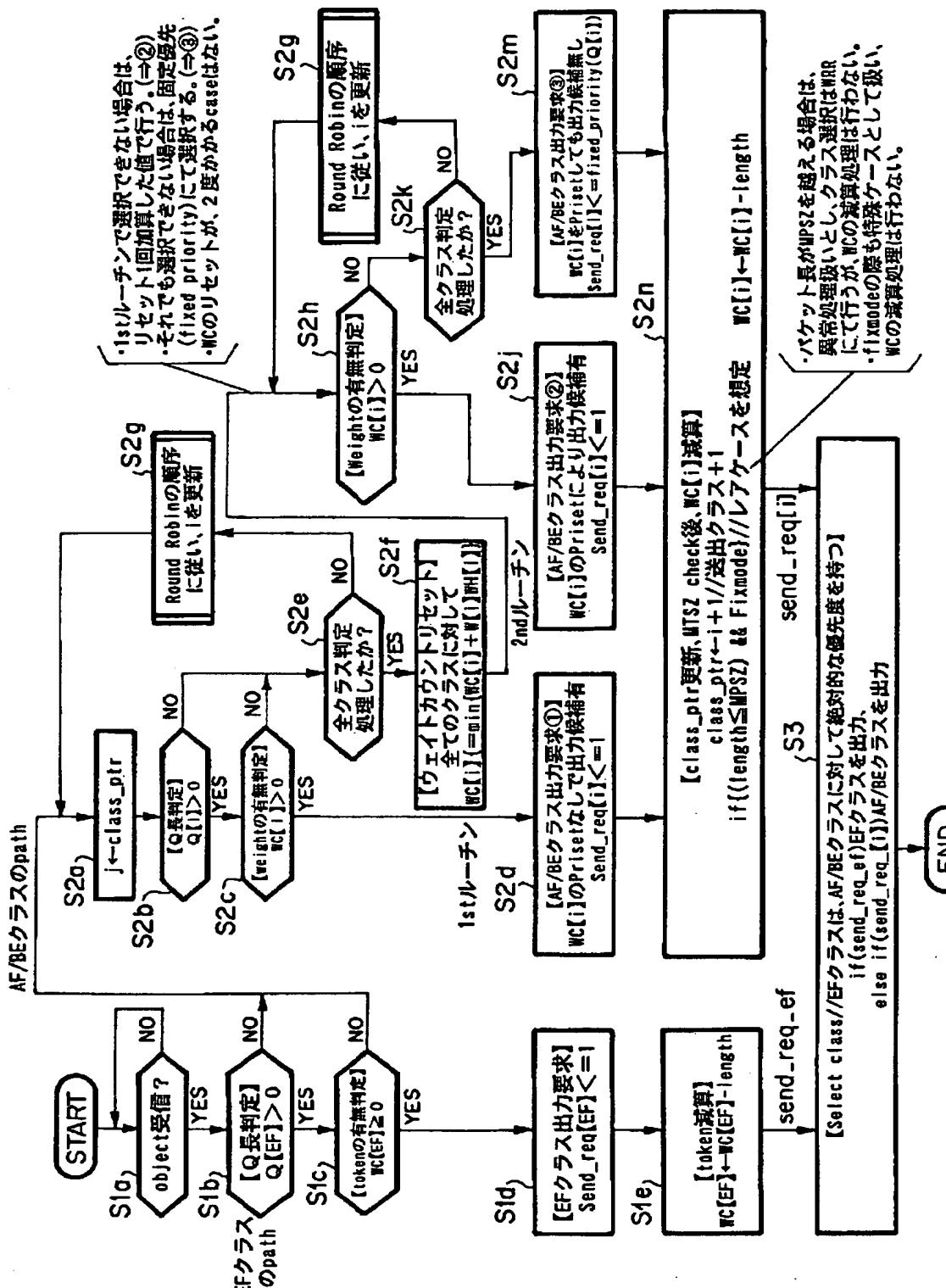
【図11】



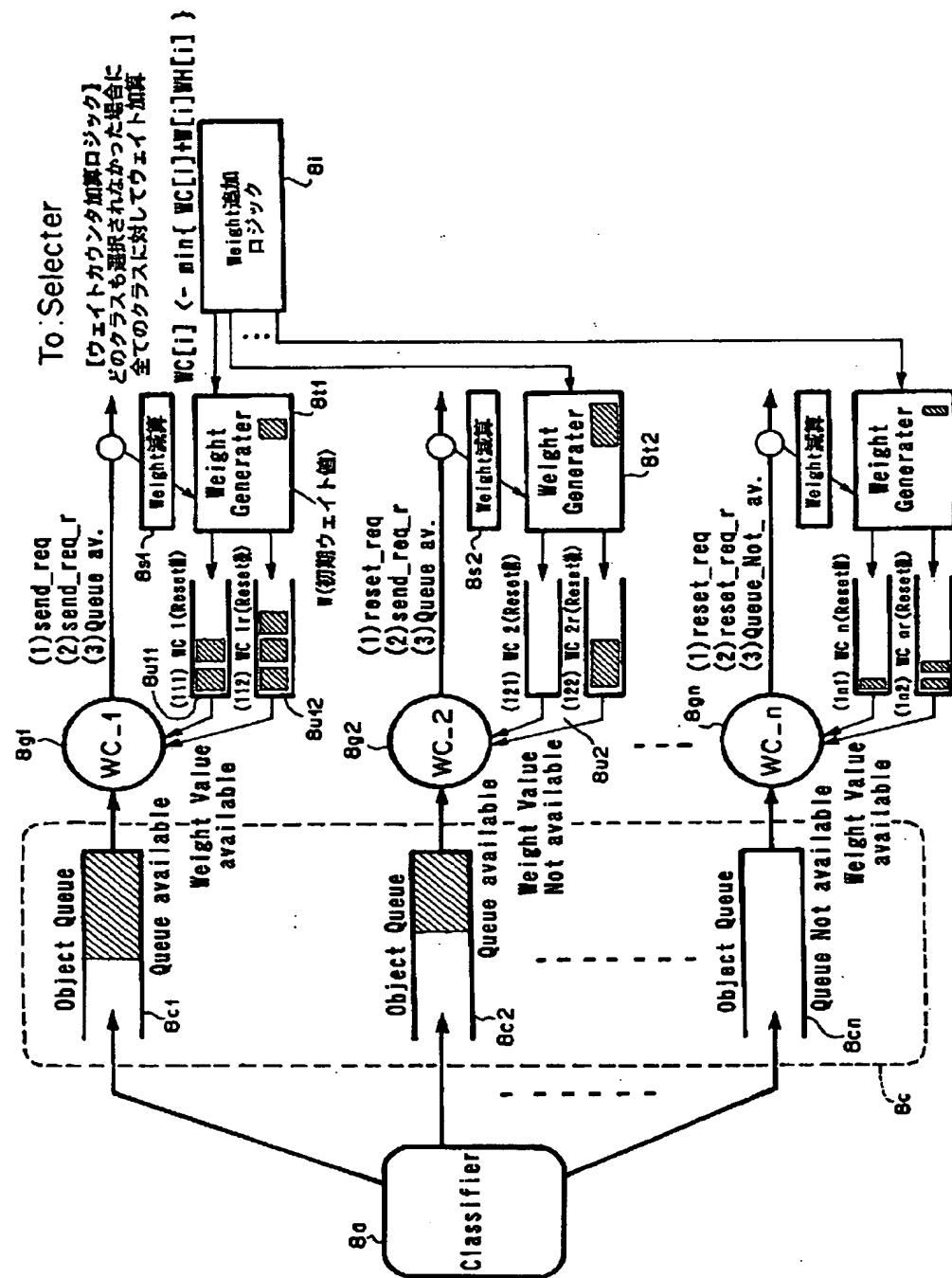
【図12】



【図13】

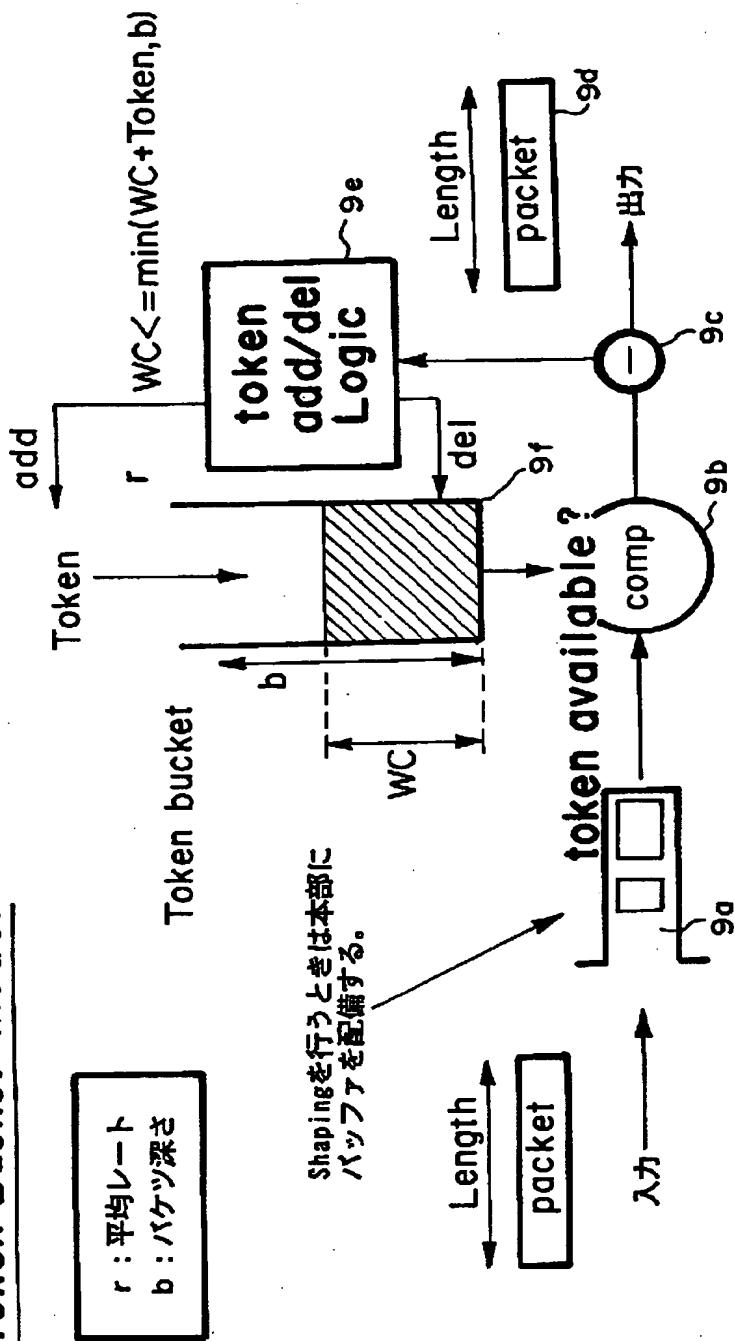


【図14】

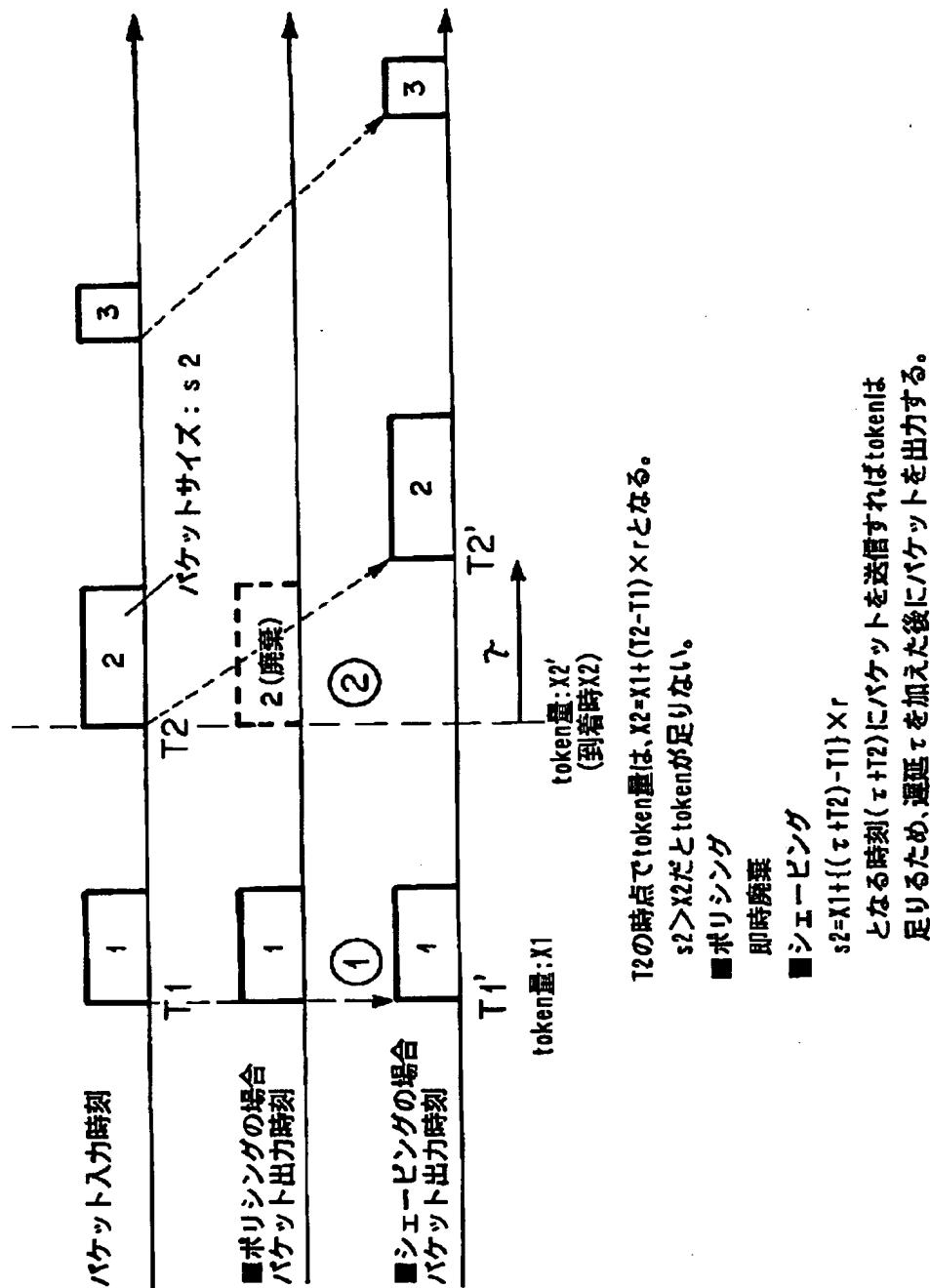


【図15】

Token Bucket Model



【図16】



【書類名】 要約書

【要約】

【課題】 パラメータの変更のみで柔軟に対応可能であり、汎用性に優れたマルチレイヤクラス識別通信装置を提供することを課題とする。

【解決手段】 装置入力インターフェースにおいて、受信したIPパケットのヘッダ情報（IPヘッダに代表されるレイヤ3情報、およびその上位レイヤに相当するTCP／UDPヘッダに代表されるレイヤ4情報の組合せ）から装置内におけるクラス識別子を解決し、個々のIPパケットフローを装置内で取り扱うサービス品質(Internet Protocol Quality of Service: IP-QoS)コードを割当てる手段を有し、上記IP-QoSコード割り当て（トラヒッククラスのmapping）に関し、IPヘッダおよび、TCPヘッダ内の複数フィールドの任意の組み合わせで優先トラヒック規定できる手段を有することを特徴とする。

【選択図】 図1

出願人履歴情報

識別番号 [000004237]

1. 変更年月日 1990年 8月29日

[変更理由] 新規登録

住 所 東京都港区芝五丁目7番1号

氏 名 日本電気株式会社